

SEARCH REQUEST FORM**Scientific and Technical Information Center**

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY**Type of Search****Vendors and cost where applicable**

Searcher: Baker, E 4994 NA Sequence (#) _____ STN _____
 Searcher Phone #: _____ AA Sequence (#) _____ Dialog _____
 Searcher Location: _____ Structure (#) _____ Questel/Orbit _____
 Date Searcher Picked Up: _____ Bibliographic _____ Dr.Link _____
 Date Completed: 01-23-02 Litigation _____ Lexis/Nexis _____
 Searcher Prep & Review Time: 3 Fulltext _____ Sequence Systems _____
 Clerical Prep Time: _____ Patent Family _____ WWW/Internet _____
 Online Time: 20 Other _____ Other (specify) CGN

piegler
09/18/86034
Seq. ID 2 w/ Interf

GenCode version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:18:25 ; Search time 96.29 Seconds
(without alignments) 801.627 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHRAGLLGLCARANSVRMA.....DASYITGETVVVGGSPTSL 278

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main : *
1: /cgn2.6/ptodata/2/paa/US06_COMB.pcp.*
2: /cgn2.6/ptodata/2/paa/US06_COMB.pcp.*
3: /cgn2.6/ptodata/2/paa/US07_COMB.pcp.*
4: /cgn2.6/ptodata/2/paa/US08_COMB.pcp.*
5: /cgn2.6/ptodata/2/paa/US081_COMB.pcp.*
6: /cgn2.6/ptodata/2/paa/US082_COMB.pcp.*
7: /cgn2.6/ptodata/2/paa/US083_COMB.pcp.*
8: /cgn2.6/ptodata/2/paa/US084_COMB.pcp.*
9: /cgn2.6/ptodata/2/paa/US085_COMB.pcp.*
10: /cgn2.6/ptodata/2/paa/US086_COMB.pcp.*
11: /cgn2.6/ptodata/2/paa/US087_COMB.pcp.*
12: /cgn2.6/ptodata/2/paa/US088_COMB.pcp.*
13: /cgn2.6/ptodata/2/paa/US089_COMB.pcp.*
14: /cgn2.6/ptodata/2/paa/US090_COMB.pcp.*
15: /cgn2.6/ptodata/2/paa/US091_COMB.pcp.*
16: /cgn2.6/ptodata/2/paa/US092_COMB.pcp.*
17: /cgn2.6/ptodata/2/paa/US093_COMB.pcp.*
18: /cgn2.6/ptodata/2/paa/US094_COMB.pcp.*
19: /cgn2.6/ptodata/2/paa/US095_COMB.pcp.*
20: /cgn2.6/ptodata/2/paa/US096_COMB.pcp.*
21: /cgn2.6/ptodata/2/paa/US097_COMB.pcp.*
22: /cgn2.6/ptodata/2/paa/US098_COMB.pcp.*
23: /cgn2.6/ptodata/2/paa/US099_COMB.pcp.*
24: /cgn2.6/ptodata/2/paa/US060_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1399	100.0	278	22	US-09-866-034-2
2	1399	100.0	278	23	Sequence 2, Appli
3	1353	96.7	518	1	Sequence 60, Appli
4	1327	94.9	278	1	Sequence 15496, A
5	1327	94.9	278	15	PCT-US01-14827-15496
6	1229.5	87.9	477	1	Sequence 1, Appli
7	1229.5	87.9	477	1	Sequence 1, Appli
8	1229.5	87.9	477	1	Sequence 7140, Ap
9	1210	86.5	268	1	Sequence 9680, Ap
					Sequence 15494, A
					Sequence 10624, A

10	1196	85.5	244	24	US-60-323-349-175	Sequence 175, App
11	1196	85.5	333	1	PCT-US01-08656-7142	Sequence 7142, Ap
12	1161	83.0	440	1	PCT-US01-14827-9686	Sequence 9686, Ap
13	906	64.8	193	1	PCT-US01-01239-1013	Sequence 1013, Ap
14	906	64.8	193	1	PCT-US01-01308-456	Sequence 456, App
15	906	64.8	193	1	PCT-US01-01349-673	Sequence 673, App
16	906	64.8	193	21	US-09-764-849-456	Sequence 456, App
17	906	64.8	193	21	US-09-764-849-456	Sequence 673, App
18	906	64.8	193	21	US-09-764-902-1013	Sequence 1013, Ap
19	824.5	58.9	280	1	PCT-US99-16164-5	Sequence 5, Appli
20	824.5	58.9	280	15	US-09-116-750-5	Sequence 5, Appli
21	824.5	58.9	280	24	US-60-277-380-17	Sequence 17, Appl
22	809.5	57.9	205	1	PCT-US01-14827-9684	Sequence 9684, Ap
23	797	57.0	277	24	US-60-245-241-126	Sequence 126, App
24	677	48.4	136	24	US-60-245-241-127	Sequence 127, App
25	664	47.5	317	24	US-60-167-217-22643	Sequence 22643, A
26	664	47.5	317	24	US-60-173-464-18416	Sequence 18416, A
27	664	47.5	317	24	US-60-191-637-22451	Sequence 22451, A
28	664	47.5	317	24	US-60-191-681-17735	Sequence 17735, A
29	649.5	46.4	183	1	PCT-US00-05881-530	Sequence 530, App
30	649.5	46.4	183	23	US-09-925-238-530	Sequence 530, App
31	620.5	44.4	276	24	US-60-324-109-26991	Sequence 26991, A
32	618.5	44.2	253	24	US-60-324-109-25978	Sequence 25978, A
33	616.5	44.1	265	24	US-60-324-109-16907	Sequence 16907, A
34	588.5	42.1	255	19	US-09-595-329A-1555	Sequence 1555, Ap
35	588.5	42.1	277	19	US-09-595-329A-1554	Sequence 1554, Ap
36	586.5	41.9	263	24	US-60-324-109-22645	Sequence 22645, A
37	468	33.5	221	21	US-09-760-475-1902	Sequence 1902, Ap
38	428	30.6	103	1	PCT-US00-26524B-5344	Sequence 5344, Ap
39	409	29.2	256	16	US-09-252-991A-30474	Sequence 30474, A
40	404	28.9	248	17	US-09-305-390-5	Sequence 5, Appli
41	404	28.9	248	17	US-09-305-390-5	Sequence 5, Appli
42	395.5	28.3	249	23	US-09-902-540-14561	Sequence 14561, A
43	373	26.7	249	15	US-09-134-001C-4825	Sequence 4825, Ap
44	373	26.7	249	18	US-09-450-969-5675	Sequence 5675, Ap
45	373	26.7	287	24	US-60-324-109-29150	Sequence 29150, A

ALIGNMENTS

RESULT 1
US-09-866-034-2
Sequence 2, Application US/09866034

GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2930R1C1
CURRENT APPLICATION NUMBER: US/09/866,034
CURRENT FILING DATE: 2001-05-25
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 2
LENGTH: 278
TYPE: PRT
ORGANISM: Homo sapiens
US-09-866-034-2

1	PRIOR APPLICATION NUMBER: 60/0979737
2	PRIOR FILING DATE: 1998-08-26
3	PRIOR APPLICATION NUMBER: 60/103396
4	PRIOR FILING DATE: 1998-10-07
5	PRIOR APPLICATION NUMBER: 60/108867
6	PRIOR FILING DATE: 1998-11-17
7	PRIOR APPLICATION NUMBER: 60/112851
8	PRIOR FILING DATE: 1998-12-16
9	PRIOR APPLICATION NUMBER: 60/119965
10	PRIOR FILING DATE: 1999-02-12
11	PRIOR APPLICATION NUMBER: 60/123972
12	PRIOR FILING DATE: 1999-03-11
13	PRIOR APPLICATION NUMBER: 60/133459
14	PRIOR FILING DATE: 1999-05-11
15	PRIOR APPLICATION NUMBER: 60/140650
16	PRIOR FILING DATE: 1999-06-22
17	PRIOR APPLICATION NUMBER: 60/140653
18	PRIOR FILING DATE: 1999-06-22
19	PRIOR APPLICATION NUMBER: 60/144758
20	PRIOR FILING DATE: 1999-07-20
21	PRIOR APPLICATION NUMBER: 60/145698
22	PRIOR FILING DATE: 1999-07-26
23	PRIOR APPLICATION NUMBER: 60/146222
24	PRIOR FILING DATE: 1999-07-28
25	PRIOR APPLICATION NUMBER: 60/149395
26	PRIOR FILING DATE: 1999-08-17
27	PRIOR APPLICATION NUMBER: 60/151689
28	PRIOR FILING DATE: 1999-08-31
29	PRIOR APPLICATION NUMBER: 08/625328
30	PRIOR FILING DATE: 1996-04-01
31	PRIOR APPLICATION NUMBER: 08/710802
32	PRIOR FILING DATE: 1996-09-23
33	PRIOR APPLICATION NUMBER: 08/800699
34	PRIOR FILING DATE: 1997-02-14
35	PRIOR APPLICATION NUMBER: 08/828683
36	PRIOR FILING DATE: 1997-03-31
37	PRIOR APPLICATION NUMBER: 08/829270
38	PRIOR FILING DATE: 1997-03-31
39	PRIOR APPLICATION NUMBER: 08/928069
40	PRIOR FILING DATE: 1997-09-11
41	PRIOR APPLICATION NUMBER: 08/934494
42	PRIOR FILING DATE: 1997-09-19
43	PRIOR APPLICATION NUMBER: 09/143068
44	PRIOR FILING DATE: 1998-08-28
45	PRIOR APPLICATION NUMBER: 09/143707
46	PRIOR FILING DATE: 1998-08-28
47	PRIOR APPLICATION NUMBER: 09/151889
48	PRIOR FILING DATE: 1998-09-11
49	PRIOR APPLICATION NUMBER: 09/169104
50	PRIOR FILING DATE: 1998-10-09
51	PRIOR APPLICATION NUMBER: 09/202089
52	PRIOR FILING DATE: 1998-12-08
53	PRIOR APPLICATION NUMBER: 09/254311
54	PRIOR FILING DATE: 1999-03-03
55	PRIOR APPLICATION NUMBER: 09/304003
56	PRIOR FILING DATE: 1999-04-30
57	PRIOR APPLICATION NUMBER: 09/380137
58	PRIOR FILING DATE: 1999-08-25
59	PRIOR APPLICATION NUMBER: 09/380138
60	PRIOR FILING DATE: 1999-08-25
61	PRIOR APPLICATION NUMBER: 09/380139
62	PRIOR FILING DATE: 1999-08-25
63	PRIOR APPLICATION NUMBER: 09/403297
64	PRIOR FILING DATE: 1999-10-18
65	PRIOR APPLICATION NUMBER: 09/423844
66	PRIOR FILING DATE: 1999-11-12
67	PRIOR APPLICATION NUMBER: 09/511133
68	PRIOR FILING DATE: 2000-02-23
69	PRIOR APPLICATION NUMBER: 09/511631
70	PRIOR FILING DATE: 2000-02-23
71	PRIOR APPLICATION NUMBER: 09/664610
72	PRIOR FILING DATE: 2000-09-18
73	PRIOR APPLICATION NUMBER: 09/665350

```

: PRIOR APPLICATION NUMBER: PCT/US01/21066
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: PCT/US01/21735
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 258
: SEQ ID NO 60
: LENGTH: 278

Query Match      100.0%; Score 1399; DB:23; Length 278;
Best Local Similarity 100.0%; Pred. No. 9.6e-138;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHKAGLLGLCARAWNSVRWASSGTMTRRDP LANKVALVTASTD GIGFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRWASSGTMTRRDP LANKVALVTASTD GIGFAIARRLAQDGAHV 60

QY 61 VSSRQONVDQAVATLQGEGLSVTGVCHVKAEDRERLVATAVKLHG GIDILVSNAAVN 120
Db 61 VSSRQONVDQAVATLQGEGLSVTGVCHVKAEDRERLVATAVKLHG GIDILVSNAAVN 120

QY 121 PFSGIMDVTEEWBDKTLDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFSGIMDVTEEWBDKTLDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180

QY 181 PYNVSKTALLGLTKT LAIELAPNRINVNCLAPGLIKTFSRMLWMDKEEESMKETLRLR 240
Db 181 PYNVSKTALLGLTKT LAIELAPNRINVNCLAPGLIKTFSRMLWMDKEEESMKETLRLR 240

QY 241 RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278
Db 241 RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278

RESULT 3
PCT-US01-14827-15496
; Sequence 15496, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 15496
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (35)...(53)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by eMATRIX, accession number PR00081A, p-value=6.22e-
; OTHER INFORMATION: 13, raw score of 10.53
; NAME/KEY: DOMAIN
; LOCATION: (34)...(220)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
; OTHER INFORMATION: accession name adh_short, E-value=2.7e-60, Pfam score of 213.7
PCT-US01-14827-15496

Query Match      96.7%; Score 1353; DB 1; Length 518;
Best Local Similarity 97.1%; Pred. No. 1.7e-132;
Matches 271; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

QY 1 MHKAGLLGLCARAWNSVRWASSGTMTRRDP LANKVALVTASTD GIGFAIARRLAQDGAHV 60
Db 2 MHKAGLLGLCARAWNSVRWASSGTMTRRDP LANKVALVTASTD GIGFAIARRLAQDGAHV 61

QY 61 VSSRQONVDQAVATLQGEGLSVTGVCHVKAEDRERLVATAVKLHG GIDILVSNAAVN 120

```

Db 62 VSSRKQNVDOAVATLOGEGLSVTGTCHVCKAEDRRLVATAVKLHGGIDILVSNAAVN 121
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 122 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 181
Qy 181 PYNVSKTALLGLTKTLATLAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 182 PYNVSKTVLLGLTKTLATLAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 241
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTV-----VGGG 273
Db 242 RLGPEDCAGIVSFLCSEDASYITGETVTVNLSVMTGGG 280

RESULT 4

PCT-US99-16164-1
; Sequence 1, Application PC/TUS9916164
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: SCAD-RELATED MOLECULES
; FILE REFERENCE: PF-0559 PCT
; CURRENT APPLICATION NUMBER: PCT/US99/16164
; CURRENT FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/116,750; Unassigned
; EARLIER FILING DATE: 1998-07-16; 1998-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 278
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 1240869
; PUBLICATION INFORMATION:
PCT-US99-16164-1

Query Match 94.9%; Score 1327; DB 1; Length 278;
Best Local Similarity 96.4%; Pred. No. 3.5e-130;
Matches 268; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MHKAGLLGLCARAWNSYRMASGTRRDPLANKVALYVASTDGGIGFAIARRLAQDRAHV 60
Db 1 MHMARLLGLCAWARKSVRMASSTRTRRDPLTNKVALYVASTDGGIGFAIARRLAQDRAHV 60
Qy 61 VSSRKQNVDOAVATLOGEGLSVTGTCHVCKAEDRRLVATAVKLHGGIDILVSNAAVN 120
Db 61 VSSRKQNVDOAVATLOGEGLSVTGTCHVCKAEDRRLVATAVKLHGGIDILVSNAAVN 120
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTLATLAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 181 PYNVSKTALLGLTNNTLAIELAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 240
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
Db 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278

RESULT 5

US-09-116-750-1
; Sequence 1, Application US/09116750
; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Azimzai, Yalda
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: SCAD-RELATED MOLECULES
; FILE REFERENCE: PF-0559 US
; CURRENT APPLICATION NUMBER: US/09/116,750
; CURRENT FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 278
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 1240869
US-09-116-750-1

Query Match 94.9%; Score 1327; DB 15; Length 278;
Best Local Similarity 96.4%; Pred. No. 3.5e-130;
Matches 268; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MHKAGLLGLCARAWNSYRMASGTRRDPLANKVALYVASTDGGIGFAIARRLAQDRAHV 60
Db 1 MHMARLLGLCAWARKSVRMASSTRTRRDPLTNKVALYVASTDGGIGFAIARRLAQDRAHV 60
Qy 61 VSSRKQNVDOAVATLOGEGLSVTGTCHVCKAEDRRLVATAVKLHGGIDILVSNAAVN 120
Db 61 VSSRKQNVDOAVATLOGEGLSVTGTCHVCKAEDRRLVATAVKLHGGIDILVSNAAVN 120
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTLATLAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 181 PYNVSKTALLGLTNNTLAIELAPRNRVNCPLAPGLIKTSFRLMWMDEKESMKETLRIR 240
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
Db 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278

RESULT 6

PCT-US01-08656-7140
; Sequence 7140, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7140
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (34)...(52)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by eMATRIX, accession number PR00081A, p-value=6.22
; OTHER INFORMATION: 13, raw score of 10.53
; NAME/KEY: DOMAIN
; LOCATION: (33)...(218)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,

OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-08656-7140

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Qy 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260

RESULT 7

PCT-US01-14827-9680
Sequence 9680, Application PC/TUS0114827
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-104
CURRENT APPLICATION NUMBER: PCT/US01/14827
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 16102
SOFTWARE: Custom
SEQ ID NO 9680
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (34)..(52)
OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
OTHER INFORMATION: Identified by eMATRIX, accession number PR00081A, p-value=6.226e-13, raw score of 10.53
NAME/KEY: DOMAIN
LOCATION: (33)..(218)
OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam.
OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-14827-9680

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Qy 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180

Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
RESULT 8
PCT-US01-14827-15494
Sequence 15494, Application PC/TUS0114827
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-104
CURRENT APPLICATION NUMBER: PCT/US01/14827
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 16102
SOFTWARE: Custom
SEQ ID NO 15494
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (34)..(52)
OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
OTHER INFORMATION: Identified by eMATRIX, accession number PR00081A, p-value=6.226e-13, raw score of 10.53
NAME/KEY: DOMAIN
LOCATION: (33)..(218)
OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam.
OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-14827-15494

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Qy 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLQEGSLVTVGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTKTALTELAPRNIRVNCCLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260

RESULT 9

PCT-US01-08656-10624
Sequence 10624, Application PC/TUS0108656
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066

```

; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-16
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10624
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (151)..(189)
; OTHER INFORMATION: Short-chain dehydrogenases/reductases family proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00061B, p-value=2.500e-
; OTHER INFORMATION: 16, raw score of 25.79
; NAME/KEY: DOMAIN
; LOCATION: (21)..(216)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
; OTHER INFORMATION: accession name adh_short, E-value=6.6e-27, Pfam score of 102.8
; NAME/KEY: misc_feature
; LOCATION: (1)..(268)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08656-10624

Query Match      86.5%; Score 1210; DB1; Length 268;
Best Local Similarity 92.9%; Pred. No. 6.5e-118;
Matches 249; Conservative 1; Mismatches 16; Indels 2; Gaps 2;

QY 13 AINSVRMASSGTRDRPLANKVALVTASTDGIQFAIARRLAQDGAHVVSRRKQONVDQA 72
Db 1 ARKSVMASSRMRDRPLTNKVALVTASTDGIQFAIARRLAQDGAHVVSRRKQONVDQA 60
QY 73 VATLQEGLSVTGTVCHVKGAEADR-ERLVATAVKLHGIDILVSNAAVNPFGSIMDYTE 131
Db 61 VATLQEGLSVTGTVCHVKGAEADRGAAWPPAVKLHGIDILVSNAAVNPFGSIMDYTE 120
QY 132 EVMDKTLIDINVKAP-ALMTKAVVPVMEKRGGSVVIVSSIAAFSPGFSFYNYSKATALL 190
Db 121 EVMDKTLIDINVKGPXKPMXKAVVPEMEKRGGSVVIVSSIAAFSPGFSFYNYSKATALL 180
QY 191 GLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIRRLGPEPCAG 250
Db 181 GLAQTLPTELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIRRLGPEPCAG 240
QY 251 IVSFLCSEDASYINGEVVVGSGTSPRL 278
Db 241 IVSFLCSEDASYINGEVVVGSGTSPRL 268

RESULT 10
US-60-323-349-175
; Sequence 175, Application US/60323349
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Asundi, Vinod
; APPLICANT: Goodrich, Ryle W
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Weinman, Tom
; APPLICANT: Weng, Gezhong
; APPLICANT: Zhou, Ping
; APPLICANT: Drmanac, Radomir
; TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides
; FILE REFERENCE: 808
; CURRENT APPLICATION NUMBER: US/60/323,349
```

```

; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/04927
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/04941
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US01/08656
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/14827
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: pf_genes Version 5.0
; SEQ ID NO 175
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-323-349-175
```

```

Query Match      85.5%; Score 1196; DB 24; Length 244;
Best Local Similarity 87.8%; Pred. No. 1.7e-116;
Matches 244; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 MHKAGLLGLCARAWNSYRMASSGTRDRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSYRMASSGTRDRDPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
QY 61 VSSRKQONVDQAVATLOGEGLSVTGTVCHVKGAEADRELVTATVKLHGIDILVSNAAVN 120
Db 61 VSSRKQONVDQAVATLOGEGLSVTGTVCHVKGAEADRELVTATVKLHGIDILVSNAAVN 102
QY 121 PFGSIMDVTVEVMDKTLIDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPGFS 180
Db 103 -----TLDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPGFS 146
QY 181 PYNVSKTALLGLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIR 240
Db 147 PYNVSKTALLGLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIR 206
```

```
QY 241 RLGEPCAGIVSFLCSEDASYITGETVVVGGTSPRL 278
|||||
Db 207 RLGEPCAGIVSFLCSEDASYITGETVVVGGTSPRL 244

RESULT 11
PCT-US01-08656-7142
; Sequence 7142, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7142
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (35)...(53)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by EMATRIX, accession number PR00081A, p-value=6.226e-
; OTHER INFORMATION: 13, raw score of 10.53
; NAME/KEY: DOMAIN
; LOCATION: (34)...(229)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
; OTHER INFORMATION: accession name adh_short, E-value=7.7e-24, Pfam score of 92.6
; NAME/KEY: misc_feature
; LOCATION: (1)...(333)
; OTHER INFORMATION: xaa = X or * as defined in Table 2
PCT-US01-08656-7142

Query Match 85.5%; Score 1196; DB 1; Length 333;
Best Local Similarity 88.9%; Pred. No. 2,7e-116;
Matches 248; Conservative: 6; Mismatches 21; Indels 4; Gaps 4;

QY 1 MHKAGLLGLCARWNSVRMASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 60
|||
Db 2 MHKTGLGLCARWNSVRMASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 61

QY 61 VSSRKOONDQAVATLOGGLSVTGTVCVHGKAEDR-ERLVATAVKLHGGIDILVSNAAV 119
|||||
Db 62 VSSRKOONDQAVATLOGGLSVTGTVCVHGKAEDRGAAPPAVKLHGGIDILVSNAAV 121

QY 120 NPFPGSIMDVTEVWKTLDINVKAP-ALMTKAVVPEMEKRGSGSVIVSSIAAFSPSPG 178
|||||
Db 122 NPFPGSIMDVTEVWKTLDINVKGPXMTKAVVPEMEKRGSGSVIVGLSIAAPSPSPG 191

QY 179 FSPYVNSKTALLGTLKTALIELAPNIRVNCPLAPGLIKTSFSLMWMKDEESMKETLR 238
|||
Db 182 FSLYVNSKPAALLGTLKTALIELAPNIRVNCPLAPGLIKTSFSLMWMKDEESMKETLR 241

QY 239 IRL-CEPEDCAGIVSFLCSEDAS-YITGETVVVGGTSP 275
|||||
Db 242 IKVXASPEDCAGIVSFLCSEDASLSHSLGKTVVVGGTSP 280

RESULT 12
PCT-US01-14827-9686
; Sequence 9686, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104

Query Match 83.0%; Score 1161; DB 1; Length 440;
Best Local Similarity 90.0%; Pred. No. 2e-112;
Matches 242; Conservative: 1; Mismatches 18; Indels 8; Gaps 3;

QY 13 AwnSVRMASGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHVSVSSRKQNVDDQA 72
|||||
Db 1 ARKSVRMASGMRTRDPLNKNVALETASTDGIGFAIARRLAQDRAHVFVLSRKQNVDDQA 60

QY 73 VATLOGGLSVTGTVCVHGKAEDR-ERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTE 131
|||||
Db 61 VATLOGGLSVTGTVCVHGKAEDRGAAPPAVKLHGGIDILVSNAAVNPFFGSIMDVTE 120

QY 132 EWDKTLDINVKAP-ALMTKAVVPEMEKRGSGSVIVSSIAAFSPSPGFYVNSKTALL 190
|||||
Db 121 EWDKTLDINVKGPXMTKAVVPEMEKRGSGSVIVSSIAAFSPSPGFYVNSKTALL 180

QY 191 GLTKTALIELAPNIRVNCPLAPGLIKTSFSLMWMKDEESMKETLRIRRLGEPCDAG 250
|||||
Db 181 GLTKTALIELAPNIRVNCPLAPGLIKTSFSLMWMKDEESMKETLRIRRLGEPCDAG 240

QY 251 IVSFLCSEDASYITGETVV-----VGGG 273
|||||
Db 241 IVSFLCSEDASYITGETVVVNLVNFVTTGGG 269

RESULT 13
PCT-US01-01239-1013
; Sequence 1013, Application PC/TUS0101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1013
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (103)
```


OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1013

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;

QY 1 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 121
QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 138 -----LWMDKEESMKETLRIR 155

RESULT 14
PCT-US01-01308-456
Sequence 456, Application PC/TUS0101308
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ204PCT
CURRENT APPLICATION NUMBER: PCT/US01/01308
CURRENT FILING DATE: 2001-01-14
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 495
SOFTWARE: Patentin Ver 2.0
SEQ ID NO 456
LENGTH: 193
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01308-456

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;

QY 1 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 121

QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 138 -----LWMDKEESMKETLRIR 155
QY 241 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
DB 156 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193

RESULT 15
PCT-US01-01349-673
Sequence 673, Application PC/TUS0101349
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206PCT
CURRENT APPLICATION NUMBER: PCT/US01/01349
CURRENT FILING DATE: 2001-01-14
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 673
LENGTH: 193
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01349-673

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;

QY 1 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGICARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTCHVKGAEEDRLVATVAKLHGGIDILVNAVN 121
QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGSVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 138 -----LWMDKEESMKETLRIR 155
QY 241 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
DB 156 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193

Search completed: January 22, 2002, 15:21:30
Job time: 185 sec

Result No.	Query		Length	DB	ID	Description
	Score	Match				
1	618.5	44.2	253	5	US-09-708-437-59234	Sequence 59234, A
2	612.5	43.8	253	5	US-09-708-437-52797	Sequence 52797, A
3	575.5	41.1	234	5	US-09-708-427-52798	Sequence 52798, A
4	377.5	27.0	263	5	US-09-815-242-14092	Sequence 14092, A
5	368.5	26.3	284	5	US-09-815-242-10241	Sequence 10241, A
6	353	25.2	262	5	US-09-708-437-17916	Sequence 17916, A
7	352	25.2	273	5	US-09-708-427-51897	Sequence 51897, A
8	348	24.9	246	5	US-09-815-242-5461	Sequence 5461, Ap
9	348	24.9	246	5	US-09-815-242-12123	Sequence 12123, A
10	348	24.9	246	5	US-09-815-242-12803	Sequence 12803, A
11	348	24.9	246	5	US-09-815-242-13100	Sequence 13100, A
12	348	24.9	246	5	US-09-572-8101-2	Sequence 2, Appl
13	346	24.7	251	5	US-09-815-242-5689	Sequence 5689, Ap
14	346	24.7	272	5	US-09-815-242-12688	Sequence 12688, A
15	345.5	24.7	268	5	US-09-708-437-17881	Sequence 17881, A
16	345	24.7	253	5	US-09-708-437-17917	Sequence 17917, A
17	344.5	24.6	203	5	US-09-634-955B-20	Sequence 20, Appl
18	341	24.4	268	5	US-09-708-437-17920	Sequence 17920, A
19	340.5	24.3	147	5	US-09-708-437-59235	Sequence 59235, A
20	340	24.3	242	5	US-09-815-242-10974	Sequence 10974, A
21	340	24.3	243	5	US-09-815-242-13360	Sequence 13360, A
22	340	24.3	243	5	US-09-815-242-13581	Sequence 13581, A
23	338.5	24.2	146	5	US-09-708-437-52799	Sequence 52799, A
24	338.5	24.2	203	5	US-09-634-955B-26	Sequence 26, Appl
25	336.5	24.1	203	5	US-09-634-955B-17	Sequence 17, Appl
26	332.5	23.8	250	5	US-09-708-437-17882	Sequence 17882, A

```
; NAME/KEY: misc_feature
; LOCATION: 1..234
; OTHER INFORMATION: Ceres Seq. ID 1929771
US-09-708-427-52798

Query Match      41.1%; Score 575.5; DB 5; Length 234;
Best Local Similarity 50.8%; Pred. No. 5.le-45;
Matches 120; Conservative 38; Mismatches 75; Indels 3; Gaps

QY 43 GIGFAIARRLAODGAHVVSRRKQONVDOAVATLQEGLSVTGTCHVCKAEADRELVA 102
    ||| ||| ||| ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2 GIGLAIAERLGLEGAADVISRRKKKNXAVEGLRAKGITAVGAVGVSDAQORSLIE 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 103 AVKLHGIDILYSNAANVPFGSIMDVEVDKTLIDINVKAPALMTKAVVPEMEKRGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 AVKSFHIDILYSNAAANFSDVILEMKESVIDLKLDWINDVKASILIQDAAPHLRK --GS 119

QY 163 SVIVVISIAAFSPGSPFYNSKTALLGTLTKLATIELAPRINRVNCLAPGLIKTSFSRM 222
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 120 SVIISSIIAGNPQEQLTWYXVTKTALFGLTRALAGEMGP-DTRVNCIXPGFVPTRFASF 178

QY 223 LWMDKKESMKETLRIIRLGPEDCAGTVSFLCSDASYITGETVVVGGGTPTSL 278
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 179 LTENETIRELNERTKLRLGVTDMAAAAALASDDASYITAETVIATVAGVGQSRL 234
```

```

RESULT      4
US-09-815-242-14092
; Sequence 14092, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Tarwick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows version 4.0
; SEQ ID NO 14092
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14092

```

Query Match 27.0%; Score 377.5; DB 5; Length 263;
Best Local Similarity 35.3%; Pred. No. 7.1e-27;
Matches 89; Conservative 53; Mismatches 101; Indels 9; Gaps 4;
QY 30 LANKVAVLTASTDGGIFATARRLAQGAHVHVSRRKQNDVQAVATLGGELSVTGRVCH 89
| | | | : | | | | | : | : : : : | | | |
| | | | : | | | | | : | : : : : | | | |

Db 4 LTGKTALITGASQIGEGIAVFAFAGANLILLDISPE-IEKLADELGGRHRCRTAVKAD 62
QY 90 VGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSGIMDVTEVMDKTLIDINVKAPALMT 149
Db 63 VRDFASVQAARAKETGRDILVNNAGVCR-LGNFLDMSSEDRDFHIDINIKGVNVT 121
QY 150 KAVYPENKRGSGSVIVSSIAA-FSPSPGSPYNVSKTALLGTTKTLAIELAPRNIIRVN 208
Db 122 KAVLPEMIKKRGRIVNMSSVTGDMVADGETAYALSKAAIVGLTKSLAVEYAQSGIRVN 181
QY 209 CLAPGLIKTSFRLMWMDEKEE-----SMKETLRIRRLGPEPCAGIVSFLCSEDASY 262
Db 182 AICGYPVTPMAESIARQSNDDPESVLTETAKAIPLRRLADPLEVGLAFLASDESSY 241
QY 263 ITGETVVVGGGT 274
Db 242 LTGTQNVIDGS 253

RESULT 5
US-09-815-242-10241
; Sequence 10241, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10241
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10241

Query Match 26.3%; Score 368.5; DB 5; Length 285;
Best Local Similarity 32.6%; Pred. No. 5.3e-26;
Matches 88; Conservative 56; Mismatches 117; Indels 9; Gaps 4;
QY 12 RANVSVMASGMRTRDPLANKVALVYASTDGGIGFAIRLAQDGAHVYSSRRKQNVQDQ 71
Db 8 RENSVTLCINKESIKMKGLKGTALITGALQIGEGIARTFARHGANLILLDISPE-IEK 66
QY 72 AVATLQEGISVTCVCHVGAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSGIMDVTE 131
Db 67 LADELGGRHRCRTAVADVDRDPASVAAIKRAKEGREGIDILVNNAGVCR-LGSFLDMSD 125
QY 132 EVMKDITDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAA-FSPSPGSPYNVSKTALL 190

Db 126 DDRDFHIDINIKGVNVTKAVLPEMIARKDGRIVNMSSVTGDMVADPGETAVALTKAATV 185
QY 191 GLTKTALATEAPRNIIRVNCLAPGLIKTSFRLMWMDEKEE-----SMKETLRIRRLGE 244
Db 186 GLTKSLAVEYAQSGIRVNAICPGYIVRTPMASIRQSNPEDEPESVLTETAKAIPMRRLAD 245
QY 245 PEDCAGIVSFLCSEDASYITGETVVVGGGT 274
Db 246 PLEVGLAFLASDESSYLTGTQNVIDGS 275

RESULT 6
US-09-708-427-17916
; Sequence 17916, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17916
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..262
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..262
; OTHER INFORMATION: Ceres Seq. ID 1833207
US-09-708-427-17916

Query Match 25.2%; Score 353; DB 5; Length 262;
Best Local Similarity 35.0%; Pred. No. 1.2e-24;
Matches 89; Conservative 47; Mismatches 110; Indels 8; Gaps 3;
QY 24 MTRDPLANKVALVYASTDGGIGFAIRLAQDGAHVYSSRRKQNVQDGAVALTQEGLSV 83
Db 1 MDKRWSLKGMTALVTGGASGIGYAIVEELAGFGARIHVCIDISEAKLNLSLEWKKGFQV 60
QY 84 TGTVCHVGAEDRERLVAT-AVKLHGGIDILVSNAAV---NPFFGSGIMDVTEVMDKTL 139
Db 61 SGSCVDVASRPEREELMOTVSSQFDGKLNILVSNVGVIRSKP---TTEYTEDDFAFHIS 116
QY 140 INVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPSPGSPYNVSKTALLGLTKTLAIE 199
Db 117 SNVEAAVHFSQLSHPLLKASGYSIIFVSSIAGVISFDAGSIYGLTKGALIQALAKNLACE 176
QY 200 LAPRNIIRVNCLAPGLIKTSFRLMWMDEKESMKETLRIRRLGPEPCAGIVSFLCSED 259
Db 177 WAKGCIANAVAPNVINTPLSQSYLEDVSFKALLSRTPLRGVGPEPNEVASLVAELCLPA 236
QY 260 ASYITGETVVVGGG 273
Db 237 ASYITGOTICVDGG 250

RESULT 7
US-09-708-427-51897
; Sequence 51897, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364

Query Match 24.9%; Score 348; DB 5; Length 246;


```

Db      184  PGFIVSDMTDAL--SDELKEQM1PIPLARFGODTDIANTVAFSLASKAKYITGQIHVN 241
QY      272  GG 273
Db      242  GG 243

```

RESULT 12

```

US-09-572-810A-2
; Sequence 2, Application US/09572810A
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: FabG
; FILE REFERENCE: GM10192
; CURRENT APPLICATION NUMBER: US/09/572,810A
; CURRENT FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 09/238,481
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-572-810A-2

```

Query Match 24.9%; Score 348; DB 5; Length 246;
Best Local Similarity 32.6%; Pred. No. 3.2e-24;
Matches 79; Conservative 57; Mismatches 102; Indels 4; Gaps 3;

[illegible]

RESULT 13

US-09-815-242-5689

Sequence 5689, Application US/09815242

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John B.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5689
; LENGTH: 251
; TYPE: prt
; ORGANISM: Staphylococcus aureus
US-09-815-242-5689

```

Query Match	24.7%	Score 346;	DB 5;	Length 251;
Best Local Similarity	33.6%;	Pred. No. 5e-24;		
Matches 84;	Conservative	56;	Mismatches 102;	Indels 8;
				Gaps 4;

QY	30	LANKVALVTASTDGI	GATARRLAQGAH	VSSRKQNVQ	QAVATLQCEGLSV	TCVCH	89	
Db	1	LENKAVVTGASTGI	QGAIAIAQEGAV	LYAVDIAEA	-VSETVDIK	SKNGDKAKAVVD	59	
QY	90	VGKAEDRELRIVAT	AVKLHG	GIDILV	NSAANPF	FGSIMDVT	EEWDEKTL	149
Db	60	IASEQIDNFASEIR	QFGHVDV	LFNNAAGV	DNAAGRIHE	YPTD	VDYDKIMN	119
QY	150	KAVVPEMKRGGS	VIVSIAA	FSPGFS	PNVSVK	TALLGL	TALTAELAP	209
Db	120	KMLPLMWTK	-GGSIVNT	SFSQGA	ADLYRSGV	NAAKGAVIN	FTKSIA	178
QY	210	LAPGLLKT	SF-SRLMW	DEKE	-----	ESMKET	LRIRREGE	263
Db	179	IAPGTIET	PLVDKLT	SETSD	KEGKA	FENQK	WMTPLGR	238
QY	264	TGETV	VVGG	273				
Db	239	TGETIRIDG	248					

RESULT 14

```

US-09-815-242-12688
; Sequence 12688, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 12688
; LENGTH: 272
; TYPE: PRN
; ORGANISM: Staphylococcus aureus
US-0-815-242-13688

```

Query Match	Score 346	DB 5	Length 272
24.7%			

30	QY	LANKVALVASTDGTGIFATARRLAQGAHVSSRRQQNVDOAQVLPLOGEGLSVTGTVC	89
		: : : : : : : : : : : :	
4	Db	LENKVAVVGTASTGIGQASATIAAQBGAVYLAVDIAEA-VSETVDKIKSGDKARAYV	62
		: : : : : : : : : : : :	
90	QY	VGKAEDRELRVATAVKLHGIDILVNSAAVNPFFGSIIMDVTEVWDKTLIDINVKAPALMT	149
		: : : : : : : : : : : :	
63	Db	IASEQIDINFASEIRGFGHVS/LFNNAAGVDNAAGRIHEYPTDVIDYDKINVDMRGTFLMT	122
		: : : : : : : : : : : :	
150	QY	KAVVPMWRKGGSVVVSIAAFSPSGFSPYNSKTLALLGLTTLATLAPRNIIVNC	209
		: : : : : : : : : : : :	
123	Db	KWMLPLMTMK-GGSIIVTSSFGQAADLYRSGVNAAGAVINFTKSIATIEYGRDGRANA	181
		: : : : : : : : : : : :	
210	QY	LAPGLIKTSF-SRMLWMDMKKE-----ESMKETLRIIRLGEPEDCAGIYVSLCSEDASI	263
		: : : : : : : : : : : :	
182	Db	IAPGTIETPLVDKLTGTSDEKGKGFRENQKWMTPLGRLGKPEEVGKLVVFLASDESSI	241
		: : : : : : : : : : : :	
264	QY	TGETVVVGGG	273
		: : : : : : : : : : : :	
242	Db	TGETRIDGG	251
		: : : : : : : : : : : :	

RESULT 15

US-09-708-427-17881

; sequence 17881, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV, et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: THEREBY

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; CURRENT FILING DATE: 2000-11-09

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 17881

; LENGTH: 268

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: misc-feature

; LOCATION: 1..268

; OTHER INFORMATION: xaa is any amino acid

; NAME/KEY: misc-feature

; LOCATION: 1..268

; OTHER INFORMATION: Ceres Seq. ID 1833164

US-09-708-427-17881

Query Match	24.7%	Score 345.5;	DB 5;	Length 268;
Best Local Similarity	33.5%;	Pred. No. 6.1e-24;		
Matches	88;	Conservative 54;	Mismatches 108;	Indels 13; Gaps 5;

19 MASSGMRD-----PLANKVALVASTDGIGFAIRRLAQDGAHVVSRRKQNVDAVA 74
||:| || || || || | : | : | : | : | :
1 MAKAGENSROKSWSLEGMTALVTGGSKGLGEAVVEELAMLGARVHTCARDETQLQRLR 60

75 TLOGESLVTGTGVCHYCKAEDRELRVATAVKL-HGGIDILYSNAAVNPFGSIMDYTE-- 131

OY db OY

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:15:50 ; Search time 23.33 seconds
(without alignments)
882.656 Million cell updates/sec

Title: US-09-866-034-2

Perfect score: 1399

Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGGTSPSL 278

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	1399	100.0	278	21	PRO1800, a Hep27 h
2	1395	99.7	278	22	Human protein sequ
3	1327	94.9	278	21	Short chain alcoho
4	649.5	46.4	183	21	Breast and ovarian
5	588.5	42.1	254	21	Arabidopsis thalia
6	588.5	42.1	276	21	Arabidopsis thalia
7	428	30.6	103	22	Human colon cancer
8	404	28.9	248	21	Amino acid sequenc
9	370	26.4	79	21	Human secreted pro
10	353	25.2	262	21	Arabidopsis thalia
11	349	24.9	292	22	C glutamicum prote

12	349	24.9	295	22	AA179400	Corynebacterium g1
13	349	24.9	295	22	AA179401	Corynebacterium g1
14	348	24.9	246	21	AA15707	Staphylococcus aur
15	348	24.9	246	22	AA102195	S. aureus NADPH-de
16	345	24.7	253	21	AA131819	Arabidopsis thalia
17	341	24.4	243	22	AA101032	CFE 35 protein seq
18	340	24.3	243	19	AA180670	S. pneumoniae fatt
19	340	24.3	243	21	AA15706	Streptococcus pneu
20	339	24.2	254	22	AA184973	Protein with acety
21	339	24.2	306	22	AA181644	S. epidermidis ope
22	336	24.0	248	20	AA136959	Chlamydia trachoma
23	333.5	23.8	251	20	AA134891	Chlamydia pneumoni
24	333.5	23.8	262	22	AA183032	S. epidermidis ope
25	331.5	23.7	261	21	AA196271	B. subtilis glucos
26	331.5	23.7	261	21	AA154424	Amino acid sequenc
27	331.5	23.7	263	21	AA132504	Arabidopsis thalia
28	330	23.6	273	21	AA143271	Arabidopsis thalia
29	330	23.6	273	21	AA143270	Arabidopsis thalia
30	328.5	23.5	315	17	AA189323	Rape leaf beta-ket
31	328.5	23.5	315	17	AA189322	Rape seed beta-ket
32	327.5	23.4	272	21	AA110740	B. megaterium gluc
33	327.5	23.4	340	21	AA110741	H. ghilianii/B. me
34	326.5	23.3	261	13	AA127757	Glucose dehydrogen
35	326	23.3	279	21	AA149315	Arabidopsis thalia
36	325.5	23.3	261	9	AA180590	Sequence of glucos
37	325.5	23.3	261	13	AA127756	NAD affinity gluc
38	324	23.2	262	21	AA144578	Xylitol dehydrogen
39	324	23.2	262	22	AA186337	G. suboxydans DSM
40	324	23.2	271	21	AA108828	Arabidopsis thalia
41	323.5	23.1	254	21	AA132505	Arabidopsis thalia
42	323.5	23.1	264	21	AA149317	Arabidopsis thalia
43	323.5	23.1	269	21	AA149316	Arabidopsis thalia
44	323.5	23.1	318	21	AA129313	Arabidopsis thalia
45	323.5	23.1	322	21	AA129312	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AA196729

ID AA196729 standard; Protein; 278 AA.

XX

AC AA196729;

XX 26-SEP-2000 (first entry)

DT PRO1800, a Hep27 homologue.

XX

DE

XX PRO1800; Hep27; homologue; short-chain alcohol dehydrogenase; SCAD;

KW secreted protein; transmembrane protein; recombinant production;

KW gene therapy.

XX

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX W0200036102-A2.
XX 22-JUN-2000.
XX 01-DEC-1999; 99WO-US28634.
XX 16-DEC-1998; 98US-0112851.
XX 16-DEC-1998; 98US-0113145.
XX 22-DEC-1998; 98US-0113511.
XX 12-JAN-1999; 99US-0115558.
XX 12-JAN-1999; 99US-0115565.
XX 09-FEB-1999; 99US-0115733.
XX 10-FEB-1999; 99US-0119341.
XX 12-FEB-1999; 99US-0119965.
XX 02-JUN-1999; 99WO-US12252.
XX (GETH) GENENTECH INC.
XX Botstein D, Desnoyers L, Ferrara N, Fong S, Gao W, Goddard A;
PI Gurney AL, Pan J, Roy MA, Stewart TA, Tamas D, Watanabe CK;
PI Wood W;
XX WPI; 2000-431586/37.
XX N-PSDB; AAA51259.
XX Isolated nucleic acid molecule encodes a PRO polypeptide which is a
XX transmembrane polypeptide
XX Claim 12; Fig 2; 154pp; English.
XX This is PRO1800, a putative human Hep27 protein homologue, which has
XX recently been shown to have homology to known short-chain alcohol
XX dehydrogenase (SCAD) family of proteins. The invention concerns novel
XX secreted and transmembrane proteins, designated PRO polypeptides. The
XX cDNA and gene sequences are useful in the recombinant production of PRO
XX polypeptides, as a hybridization probe to screen libraries to isolate
XX cDNAs with sequence identity to PRO polypeptides or to map the gene
XX encoding the PRO polypeptides and analyzing genetic disorders. The
XX cDNA/gene can also be used to produce transgenic animals useful for the
XX development and screening of therapeutically useful reagents. They can
XX also be used in gene therapy, e.g. to replace a defective gene.
XX Sequence 278 AA;
Query Match 100.0%; Score 1399; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 4.9e-132;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHKAGLLGLCARAWNSVRVSSGMRDRPLANKVALVTASTDGTGIGFAIRRLAQDGAHV 60
Db 1 mhkagllgicarawnsvrmasgmrdrplankvalvtastdgtgigfarrrlaqdgahv 60
QY 61 VSSRQNDVQAVATLQGLSVGTGVCHVKAEDRERLVAVAKLHGGIDILVNAVN 120
Db 61 vsrkgqndvqavatlqgelsvgtgvtchvkaedrervlavaklhggidilvnaavn 120
QY 121 PFGSTMDYTEEVDKTLIDINKAPALMTKAVVPEMEKRGGSVVIVSSIAFSPSPGFS 180
Db 121 pfgstmdyteevdktlidinkapalmkavvpeemekrggsvvsviaafspspgfs 180
QY 181 PYNVSKTALLGLTKTLATLAPRNVRNCLAPGLIKTSFMRMLWMDKEESMKETLIR 240
Db 181 pynvsktallgltktlatlaprnvrnclapglktsfmrmlwmdkeesmketlir 240
QY 241 RLGEPEDCAGIVFLCSDEASVITGETVVVGGTSPSL 278
Db 241 rlgepedcagivflcsedasyitgetvvvvggtpsl 278
RESULT 2

AAB93414
ID AAB93414 standard; Protein; 278 AA.
XX
AC AAB93414;
XX
DT 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:12620.
DE
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs
XX Claim 8; SEQ ID 12620; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX Sequence 278 AA;
Query Match 99.7%; Score 1395; DB 22; Length 278;
Best Local Similarity 99.6%; Pred. No. 1.2e-131;
Matches 277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MHKAGLLGLCARAWNSVRVSSGMRDRPLANKVALVTASTDGTGIGFAIRRLAQDGAHV 60
Db 1 mhkagllgicarawnsvrmasgmrdrplankvalvtastdgtgigfarrrlaqdgahv 60

[illegible]

QY 181 PYNVSKTALLGLTKLTALAPNIRVNCPLAPGLIKTSFMRMLWMDKEESMKETLIR 240
 Db 181 pynvsktallglntlatelapnirvncplapglktsfmrmlwmdkeesmketlir 240
 QY 241 RLGEPEDCAGIVFLCSEDASYITGETWVGGGTPSRL 278
 Db 241 rlgepedcagivflcsedasyitgetvvvggtpsl 278

RESULT 4
 AAB58822
 ID AAB58822 standard; Protein; 183 AA.
 AC AAB58822;
 XX
 DT 27-MAR-2001 (first entry)
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 530.
 DE Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnary; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX Homo sapiens.
 OS
 XX
 XX
 PN W020005173-A1.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-0505881.
 XX
 XX 12-MAR-1999; 99US-012470.
 XX
 XX (HUMA-) HUMAN GENOME SCI. INC.
 PA
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-611515/58.
 DR N-PSDB; AAF21725.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases.
 XX
 XX Claim 11; Page 963-964; 1299pp; English.
 PS
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antiulcer; vulnary; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiac activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SQ Sequence 183 AA;

Query Match 46.4%; Score 649.5; DB 21; Length 183;
 Best Local Similarity 55.2%; Pred. No. 4e-57;
 Matches 148; Conservative 2; Mismatches 21; Indels 97; Gaps 3;
 QY 16 SVRWASSGTRRDPLANKVALVTASTDGGIGFATARRLAODGAHVSVSRKQONVDQAV-- 73
 Db 8 svrmassmrtrrdpltnkvalvtastdggf-----aspvvprrtgprgrqgeaaccg 61
 QY 74 ---ATLQGEGLSVTGVCHVKAEDRELRVATVKLHGDIILVSNAAVNPFFGSIIMDVT 130
 Db 62 pgggtlqgeglsvtgcxxxgkaedrervlrvatavkklhgdiilvsnaaavnpffgsimdv 121
 QY 131 EYVDKTLINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGSPYNVSKTALL 190
 Db 122 evwdk----- 127
 QY 191 GLTKTLAIELAPNIRVNCPLAPGLIKTSFMRMLWMDKEESMKETLIRRLGPEPCAG 250
 Db 128 -----lwm dkeesmketlirrlgpepcag 155
 QY 251 IVSFLCSEDASYITGETWVGGGTPSRL 278
 Db 156 ivsflcsedasyitgetvvvggtpsl 183

RESULT 5
 AAG41219
 ID AAG41219 standard; Protein; 254 AA.
 AC AAG41219;
 XX
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 51256.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 XX EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.

PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276;
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0145986.
PR	21-MAY-1999;	99US-0135253.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139222.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153378.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143342.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144085.	PR	08-OCT-1999;	99US-0158232.
PR	16-JUL-1999;	99US-0144086.	PR	12-OCT-1999;	99US-01

PR	30-APR-1999;	9905-0132048.
PR	30-APR-1999;	9905-0132407.
PR	04-MAY-1999;	9905-0132484.
PR	05-MAY-1999;	9905-0132485.
PR	06-MAY-1999;	9905-0132486.
PR	07-MAY-1999;	9905-0132487.
PR	11-MAY-1999;	9905-0132486.
PR	11-MAY-1999;	9905-0132485.
PR	14-MAY-1999;	9905-0134218.
PR	14-MAY-1999;	9905-0134219.
PR	14-MAY-1999;	9905-0134221.
PR	18-MAY-1999;	9905-0134370.
PR	18-MAY-1999;	9905-0134768.
PR	19-MAY-1999;	9905-0134941.
PR	20-MAY-1999;	9905-0135124.
PR	21-MAY-1999;	9905-0133353.
PR	24-MAY-1999;	9905-0133629.
PR	25-MAY-1999;	9905-0136021.
PR	27-MAY-1999;	9905-0136392.
PR	28-MAY-1999;	9905-0136782.
PR	01-JUN-1999;	9905-0137222.
PR	03-JUN-1999;	9905-0137528.
PR	04-JUN-1999;	9905-0137502.
PR	07-JUN-1999;	9905-0137724.
PR	08-JUN-1999;	9905-0138094.
PR	10-JUN-1999;	9905-0138540.
PR	10-JUN-1999;	9905-0138847.
PR	14-JUN-1999;	9905-0139119.
PR	16-JUN-1999;	9905-0139452.
PR	16-JUN-1999;	9905-0139453.
PR	18-JUN-1999;	9905-0139458.
PR	18-JUN-1999;	9905-0139459.
PR	18-JUN-1999;	9905-0139460.
PR	18-JUN-1999;	9905-0139461.
PR	18-JUN-1999;	9905-0139462.
PR	18-JUN-1999;	9905-0139463.
PR	18-JUN-1999;	9905-0139750.
PR	28-JUN-1999;	9905-0140693.
PR	28-JUN-1999;	9905-0140695.
PR	29-JUN-1999;	9905-0140991.
PR	30-JUN-1999;	9905-0141287.
PR	01-JUL-1999;	9905-0141842.
PR	01-JUL-1999;	9905-0142154.
PR	02-JUL-1999;	9905-0142055.
PR	06-JUL-1999;	9905-0142390.
PR	08-JUL-1999;	9905-0142803.
PR	09-JUL-1999;	9905-0142920.
PR	12-JUL-1999;	9905-0142977.
PR	13-JUL-1999;	9905-0143542.
PR	14-JUL-1999;	9905-0143624.
PR	15-JUL-1999;	9905-0144005.
PR	16-JUL-1999;	9905-0144085.
PR	16-JUL-1999;	9905-0144086.
PR	19-JUL-1999;	9905-0144325.
PR	19-JUL-1999;	9905-0144331.
PR	19-JUL-1999;	9905-0144332.
PR	19-JUL-1999;	9905-0144333.
PR	19-JUL-1999;	9905-0144334.
PR	19-JUL-1999;	9905-0144335.
PR	20-JUL-1999;	9905-0144352.
PR	20-JUL-1999;	9905-0144632.
PR	20-JUL-1999;	9905-0144884.
PR	21-JUL-1999;	9905-0144881.

RESULT	6
AAG41218	
ID	AAG41218 standard; Protein: 276 AA.
XX	
AC	AAG41218;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 51255.
KW	Protein identification; signal-transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
PN	EP1033405-A2.
PD	06-SEP-2000.
PF	25-FEB-2000; 2000EP-0301439.
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123490.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0123684.
PR	25-MAR-1999; 99US-0123664.
PR	29-MAR-1999; 99US-0125785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.

Query Match 42.1%; Score 588.5; DB 21; Length 276;
Best Local Similarity 50.6%; Pred. No 9.5e-51;
Matches 126; Conservative 35; Mismatches 85; Indels 3; Gaps

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157417.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159607.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

QY 30 LANKVALVTASTDGGFAIARRLAQDGAHVSVSRKQNVDAVATLQGEGLSVTGVCH 89
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 31 legkvaivtastgqigfiterfglegasvvsrskqanvdeavakiksgidaygivch 90
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 90 VGKAEDRERLVATAVKLHGIDILVSNAAVNPFPGSTMDYTEEWDKTLDINVKAPALMT 14
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 91 vsnaqhrrnlvektvgkygidivvcnaaanpstdpilsskeavldklweinvksilll 15
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 150 KAVPEMEKRGGGWSIVSSIAAFSPGSPGYNNVXTALLGLTKTLAIELAPRNIRVNC 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 151 qdmaphlek--gssvfitsiagfspggamamygvtkallglckaaemmap-dtrvna 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 210 LAPFLIKTSFRLMWMDEKEESMKETLRIRRGEPEDCAGIVSFLCSSEASITGETVV 26
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 208 vapgfvpthfasitgssevregleektlrlngttgdmaaaaafldsdssyitgetlv 26
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 270 VGGGTFSRL 278
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 vaggmbsrl 276
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
AAG74580
ID AAG74580 standard; Protein; 103 AA.
XX
AC AAG74580;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:5344.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX

XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.
DR N-PSDB; AAH33985.
XX
PT Nucleic acids encoding 4277 hman colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 6991-6992; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAH77788 represent human colon
CC cancer-associated nucleic acid-molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropiate p
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of p by expressing
CC inactive proteins or to supplement the patients own production of p.
CC Additionally, N may be used to produce the colon cancer-associated ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37195 to AAH37204
CC and AAH77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 103 AA;

Query Match 30.6%; Score 428; DB 22; Length 103; >
Best Local Similarity 94.6%; Pred. No. 2.8e-35;
Matches 87; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MHKAGLLGICARAWNSVRMASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 60
Dy 2 mhkagllgicarawnsvrmasgmrtrdplankvalvtastdgigfaiarrlaqdgahv 61
Qy 61 VSSRKQNVDOAVATLOGEGLSVTCTVCHVGK 92
Dy 62 vssrkqnvdaqavatlgeglsvtgcawmvr 93

RESULT 8
AA54422
ID AA54422 standard; Protein; 248-AA.
XX
AC AA54422;
XX
DT 06-APR-2000 (first entry)
XX
DE Amino acid sequence of a beta-ketoacyl-ACP reductase protein.
XX
KW Beta-ketoacyl-ACP reductase; Type II fatty acid synthetase;
KW stereoselectivity; 4-chloroacetoacetic acid ester;
KW (S)-4-halo-3-hydroxybutyric acid ester; 4-halo-acetoacetic acid ester;
KW beta-ketoacyl-acyl carrier protein reductase; acetoacetyl-CoA reductase;
KW polybeta-hydroxy fatty acid biosynthesis; optically active;
KW 4-haol-3-hydroxybutyric acid ester.
XX
OS Bacillus subtilis.
XX
XX EP955375-A2.
XX
XX 10-NOV-1999.
XX
XX 10-MAY-1999; 99EP-0109403.
XX
XX 08-MAY-1998; 98JP-0126507.
XX 21-OCT-1998; 98JP-0300178.
XX 05-APR-1999; 99JP-0098205.
XX

PA (DAIL) DAICEL CHEM IND LTD.
XX Yamamoto H;
XX
DR WPI: 2000-118183/11.
DR N-PSDB; AAZ45749.
XX
PT Producing optically pure (S)-4-halo-3-hydroxybutyric acid ester -
XX
PS Claim 6; Page 19-20; 34pp; English.
XX
CC The present sequence represents a beta-ketoacyl-ACP reductase protein
CC of Bacillus subtilis. The beta-ketoacyl-ACP reductase enzyme constitutes
CC a Type II fatty acid synthetase. The enzyme has an extremely high
CC reducing activity and stereoselectivity towards 4-chloroacetoacetic
CC acid ester. The specification describes a method for producing a
CC (S)-4-halo-3-hydroxybutyric acid ester. The method comprises
CC asymmetrically reducing 4-halo-acetoacetic acid ester or its
CC derivative with beta-ketoacyl-acyl carrier protein reductase
CC constituting Type II fatty acid synthase, or acetoacetyl-CoA
CC reductase constituting the polybeta-hydroxy fatty acid biosynthesis
CC system. The novel method is used to produce optically active
CC 4-haol-3-hydroxybutyric acid ester, with a high purity.
XX
SQ Sequence 248 AA;

Query Match 28.9%; Score 404; DB 21; Length 248;
Best Local Similarity 37.7%; Pred. No. 2.6e-32;
Matches 93; Conservative 50; Mismatches 100; Indels 4; Gaps 3;
Qy 28 DPLANKVALVTASTDGIGFAIARRLAQDGAHVVS-SRKQNVDOAVATLOGEGLSVTGT 86
Dy 2 dmlndktaivgasgrigrslalalaksganvvvnygneakanevdeiksmgrkaiav 61
Qy 87 VCHVGKAEDEERLAVATYKHLGGIDILVSNAAVNPFPGSIMDVTEVMDKTLIDINVKAPA 146
Dy 62 kadvspedvqmiketslsvfstidilvonnagitr-dnlmrmkedwdvvinolkgvf 120
Qy 147 LMTKAVPEMEKRGGSVVIVSSIAAFSPGCFSPYVSKTALLGLTKTLAIELAPRNR 206
Dy 121 nctkavtrgmkmqrsgrilnsvssivgsgnpgqanyvaakagvigtksakelasnrit 180
Qy 207 VNCLAPGLIKTSFSRMLWMDKEESMKETLIRRLGPEDCAGIVSFLCSDESAYITGE 266
Dy 181 vnaiapggfistdmtkl--akdvqdemikqiplarfgepsdvssvvtflasegarmytlgq 238
Qy 267 TVVVGGS 273
Dy 239 tllhdgg 245

RESULT 9
AAG02114
ID AAG02114 standard; Protein; 79 AA.
XX
AC AAG02114;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 6195.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX


```

PR XX 26-FEB-1999;    99US-0122487.
XX PA      (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX PS N-PSDB; AAC02120.
XX PT
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PT Claim 13; SEQ ID 6195; 71pp + CD-ROM; English.
XX CC The present sequence is a polypeptide encoded by one of a large number
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30
XX CC different tissues. EST sequences usually correspond mainly to the 3',
XX CC untranslated region (UTR) of the mRNA because they are often obtained
XX CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX CC those cases where longer cDNA sequences have been obtained, the full 5'
XX CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX CC ends and can therefore be used to obtain full length cDNAs and genomic
XX CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX CC chromosome mapping procedures. They are used to obtain upstream
XX CC regulatory sequences and to design expression and secretion vectors.
XX SS Sequence       79 AA;
XX SQ
Query Match          26.4%; Score 370; DB 21; Length 79;
Best Local Similarity 96.2%; Pred. NO. 1.3e-29;
Matches 76; Conservative 0; Mismatches 3; Indels 0; Gaps

QY 1 MHKAGLLGLCARAWNSVRMASSGWNRDPDLANKVALVTASTDGIGFAIARRLAQDGAHVV 60
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
     1 mxaxlglcrlcarawnsvrmasgwtirpdlankvalvtastdggfaiarrlaqdgahv 60

QY 61 VSSRKQQNVDAQVATLQGE 79
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
     61 vsrkrqqnvdaqvatlgge 79

RESULT 10
ID AAG31818
AA AAG31818 standard; Protein; 262 AA.
AC AC
AC AAG31818;
XX DT
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38275.
DE DE
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD
XX DP 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999;    99US-Q121825.
PR 05-MAR-1999;    99US-0123180.
PR 09-MAR-1999;    99US-Q123548.
PR 23-MAR-1999;    99US-0125788.
PR 25-MAR-1999;    99US-0126264.
PR 29-MAR-1999;    99US-0126785.
```

```

PR    13-OCT-1999;          99US-0159293.
PR    13-OCT-1999;      99US-0159294.
PR    13-OCT-1999;      99US-0159295.
PR    14-OCT-1999;      99US-0159329.
PR    14-OCT-1999;      99US-0159330.
PR    14-OCT-1999;      99US-0159331.
PR    14-OCT-1999;      99US-0159637.
PR    14-OCT-1999;      99US-0159638.
PR    18-OCT-1999;      99US-0159584.
PR    21-OCT-1999;      99US-0160741.
PR    21-OCT-1999;      99US-0160767.
PR    21-OCT-1999;      99US-0160768.
PR    21-OCT-1999;      99US-0160770.
PR    21-OCT-1999;      99US-0160814.
PR    21-OCT-1999;      99US-0160815.
PR    22-OCT-1999;      99US-0160980.
PR    22-OCT-1999;      99US-0160981.
PR    22-OCT-1999;      99US-0160989.
PR    25-OCT-1999;      99US-0161404.
PR    25-OCT-1999;      99US-0161405.
PR    25-OCT-1999;      99US-0161406.
PR    26-OCT-1999;      99US-0161359.
PR    26-OCT-1999;      99US-0161360.
PR    26-OCT-1999;      99US-0161361.
PR    28-OCT-1999;      99US-0161920.
PR    28-OCT-1999;      99US-0161992.
PR    28-OCT-1999;      99US-0161993.
PR    29-OCT-1999;      99US-0162142.

Query Match               25.2%; Score 353; DB 21; Length 262;
Best Local Similarity     35.0%; Pred. No. 3.7e-27;
Matches   89; Conservative 47; Mismatches 110; Indels 8; Gaps

QY    24 MTRRDPLANKVALVTASTDIGFAIARRLAODGAHVVSRRKOONVDQAATLQGEGLSV 83
Db    | :| | |||| ||||| || || | : |:| : |:| :| |
Dd    1 mdkrwlsgmtalvtgagsigvaiveelagfarihvcdiseakinqslsewkkqgv 60
QY    84 TGVNCHVGKRAEDRERIVAT-AVKLGHGIDILYSNAAV---NPFFGSIMDVTEEVWDKTLD 139
Db    | :| | | |||| :| | :| :| :| :| | | | | | | | | | | | | | | | |
Dd    61 sgsvcdvasrpereelmqtvsqfdgklnilvsnvgvirskp-----tteyteddfafhis 116
QY    140 INKVAPALMTKAVPPEMEKGGSGVVVISIAAFSPGFPFYNVSKTALLGLTKTLAIE 199
Db    | :| | :| | :| | :| | :| | :| | | | | | | | | | | | | | | |
Dd    117 snveayhfsqlshpllkasgygsiiifvsdiagsfdagsiygltkgaliglaknlace 176
QY    200 LAPNRIRVNCLAGLIKTFSFRMLWMDKEESMKETLRIRLGEPPDCAGIVSLCSED 259
Db    | | | | | | | | | | :| :| :| :| :| :| :| :| :| :| :| :| :| | |
QY    177 wakdgiranavapnvlntpsldysledvfakkallsrtlgrvgepnivasivalfcipa 236
Db    | | | | | :| | | | | :| :| :| :| :| :| :| :| :| :| :| :| :| | |
QY    260 ASYTGETVVVGGG 273
Db    ||||||| :| | | |

RESULT 11
AAG92082
ID AAG92082 standard; Protein; 292 AA..
XX AC
AC AAC
XT AC
DT XX
26-SEP-2001 (first entry)
DE C glutamicum protein fragment SEQ ID NO: 5836.
XX DE
KW Corynebacterium bacterium; amino acid synthesis; vitamin; saccharide;
KK organic acid synthesis.
XX OS
XX Corynebacterium glutamicum.
PN EP108790-A2.
PX PN
XX 20-JUN-2001.
```


CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 XX
 SQ Sequence 295 AA;

Query Match 24.9%; Score 349; DB 22; Length 295;
 Best Local Similarity 34.5%; Pred. No. 1.1e-26;
 Matches 86; Conservative 45; Mismatches 114; Indels 4; Gaps 2;

QY 30 LANKVALVTASTDGIQFAIRRLAQDGAHVVS--SRKQNVNDAVATLQEGSLSVGTGV 87
 Db 48 lgrkalitgdsigaavalaayaregadvaiaylpeeqadadrvlqaleetgqkafsf 107
 QY 88 CHVGKAEDERLVATAVKLHGIDILVSNAAVNPFFGSIMDVTEVMDKTLDINVKAPAL 147
 Db 108 gdlrdpeycrslvqetnalggdlvnnasrqwvpglteitdenfdqtlqnlvysfr 167
 QY 148 MTKAVPEMEKRGGSVVIVSSIAAFSPGFSFYNVSKTALLGLTKTLATELAPRIRV 207
 Db 168 vtkaaphl--kpgssliftssiqaypsetlldyamtkaalnnlsglassllgdlgrv 225
 QY 208 NCLAPGLIKTSFRLMNDKEESMKETLIRLGPEDCAGIVSFLCSDASYITGET 267
 Db 226 nsavpgpftwlpqshgqgqkiefgqhagpigrhvpelagayvflasdeasyvvet 285
 QY 268 VVGGGTPS 276
 Db 286 lgvctgtpt 294

RESULT 13
 AAB79401
 ID AAB79401 standard; Protein: 295 AA.
 XX
 AC AAB79401;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum SMP protein sequence SEQ ID NO:318.
 XX
 KW Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN W0200100844-A2.
 XX
 PD 04-JAN-2001.
 XX
 XX 23-JUN-2000; 2000WO-IB00943.
 XX
 PR 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031412.
 PR 08-JUL-1999; 99DE-1031413.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031431.
 PR 08-JUL-1999; 99DE-1031433.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031562.

PR 08-JUL-1999; 99DE-1031634.
 PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032230.
 PR 09-JUL-1999; 99US-0143208.
 PR 14-JUL-1999; 99DE-1032924.
 PR 14-JUL-1999; 99DE-1032973.
 PR 14-JUL-1999; 99DE-1033005.
 PR 27-AUG-1999; 99DE-1040765.
 PR 31-AUG-1999; 99US-0151572.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042123.
 PR 03-SEP-1999; 99DE-1042125.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;
 XX
 DR WPI: 2001-061975/07.
 DR N-PSDB; AAF71518.

XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes .
 XX

PS Claim 20; Page 587-588; 1246pp; English.

XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in
 CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
 CC energy production. The C. glutamicum SMP gene can be used in vectors
 CC (II) for expression in host cells and production or modulation of
 CC production of fine chemicals, such as, an organic acid, a proteinogenic
 CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 CC (III) encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 XX

SQ Sequence 295 AA;

Query Match 24.9%; Score 349; DB 22; Length 295;
 Best Local Similarity 34.5%; Pred. No. 1.1e-26;
 Matches 86; Conservative 45; Mismatches 114; Indels 4; Gaps 2;

QY 30 LANKVALVTASTDGIQFAIRRLAQDGAHVVS--SRKQNVNDAVATLQEGSLSVGTGV 87
 Db 48 lgrkalitgdsigaavalaayaregadvaiaylpeeqadadrvlqaleetgqkafsf 107
 QY 88 CHVGKAEDERLVATAVKLHGIDILVSNAAVNPFFGSIMDVTEVMDKTLDINVKAPAL 147
 Db 108 gdlrdpeycrslvqetnalggdlvnnasrqwvpglteitdenfdqtlqnlvysfr 167
 QY 148 MTKAVPEMEKRGGSVVIVSSIAAFSPGFSFYNVSKTALLGLTKTLATELAPRIRV 207
 Db 168 vtkaaphl--kpgssliftssiqaypsetlldyamtkaalnnlsglassllgdlgrv 225
 QY 208 NCLAPGLIKTSFRLMNDKEESMKETLIRLGPEDCAGIVSFLCSDASYITGET 267
 Db 226 nsavpgpftwlpqshgqgqkiefgqhagpigrhvpelagayvflasdeasyvvet 285

Db 236 nsavpgftwplqpsghgqgkiefgqhagpigraghvlagayvflasdeasyvvgat 285

Qy 268 VVVGGTFS 276
 Db 286 lvgtgtpt 294

RESULT 14
 AAB15707
 ID AAB15707 standard; Protein; 246 AA.
 AC AAB15707;
 XX
 DT 07-DEC-2000 (first entry)
 XX
 DE Staphylococcus aureus fabg polypeptide.
 XX
 KW Staphylococcus aureus; fabg; 3-ketoacyl-ACP reductase; antibacterial;
 KW cytosatic; antiulcer; cancer; gastric ulcer; gastritis;
 KW Helicobacter pylori infection; microbial infection.
 XX
 OS Staphylococcus aureus.
 XX
 XX WOZ00044773-Al.
 PN
 XX
 PD 03-AUG-2000.
 XX
 PF 19-JAN-2000; 2000WO-US01196.
 XX
 PR 28-JAN-1999; 99US-0238481.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Huang J, Mcdevitt D;
 XX
 XX WPI; 2000-482968/42.
 DR
 DR N-PSDB; AAA74685.
 XX

Fabg polypeptide, isolated from Staphylococcus aureus, is used to treat microbial diseases, identify agonists and antagonists for treating microbial infections and to detect diseases associated with microbial infections.

Claim 1; Page 3; 37pp; English.

The present sequence is a fabg (3-ketoacyl-ACP reductase) polypeptide from Staphylococcus aureus. A full length fabg gene was isolated from an S. aureus WCUH 29 DNA library. fabg polynucleotides and polypeptides are used for detection and treatment of microbial diseases. They may also be used to identify antagonists and agonists which can then be used to treat microbial diseases. Compounds that interfere with the initial physical interaction between a pathogen and a host have been identified. The compounds are able to prevent the adhesion of bacteria to mammalian extracellular proteins in wounds, prevent adhesion between mammalian extracellular proteins and bacterial fabg proteins which mediate tissue damage and/or to block normal progression of pathogenesis in infections mediated by implantation of in-dwelling devices or other surgical techniques. The fabg polypeptides, and polynucleotides, antagonists and agonists are especially useful in the treatment of Helicobacter pylori infection. They may be used to decrease H. pylori-induced cancers and to prevent, inhibit and/or cure gastric ulcers and gastritis.

SQ Sequence 246 AA;

Query Match 24.98; Score 348; DB 21; Length 246;
 Best Local Similarity 32.68; Pseq No. 11e-26;
 Matches 79; Conservative 57; Mismatches 102; Indels 4; Gaps 3;

Qy 33 KVALVTSTGIGFAIRLARLDGAAHVVS-SRKQNNVDQAVATLQEGSLVSTGTVCHVG 91
 Db 5 ksalvtgstrigrsialqiaeeynnvavnyagskkaevveikakgvdslaiqanva 64

Qy 92 KAEDRELAVATVAKLHGIDILVSNAAVNPFFGSSIMDVTEEVWDKTLDINVKAPALMTKA 151
 Db 65 dadevkanikvsvqfsgldvlvnnagitr-dnllrmkqegwddvidtnlkgvfnclqk 123

Qy 152 VVPEMEKRGGSVVIVSSIAAFSPSPYVNVSTALLGLTKTLAIELAPNIRVNCLA 211
 Db 124 atpdmirgrsgainlssvvavgnpgqanyvatkgavigtksaarelasrgitvnava 183

Qy 212 PGLIKTSFSLMMDKEESMKETLRRRLGEPEDCAGIVSFLCSESDASYITGFTVVVG 271
 Db 184 pgfivsdmtal--sdelkeqmltqiariqgtdiantvaflaskakytgqtihn 241

Qy 272 GG 273
 Db 242 gg 243

RESULT 15
 AAE02195
 ID AAE02195 standard; Protein; 246 AA.
 AC AAE02195;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE S. aureus NADPH-dependent beta-ketoacyl-ACP reductase (FabG).
 XX
 KW FabG; high throughput method; fatty acid biosynthesis; therapy;
 KW bacterial enzyme; biological agent screening; otitis media; empyema;
 KW bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;
 KW infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;
 KW intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;
 KW conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;
 KW cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;
 KW impetigo; folliculitis; wound infection; bacterial myositis;
 KW septic arthritis; osteomyelitis; beta-ketoacyl-ACP reductase;
 KW acyl carrier protein.
 XX
 XX Staphylococcus aureus 'WCUH 29'.
 OS
 XX WOZ00130988-Al.
 PN
 XX
 PD 03-MAY-2001.
 XX
 XX 26-OCT-2000; 2000WO-US29451.
 PF
 XX
 PR 27-OCT-1999; 99US-0161775.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Dewolf W, Kallender H, Lonsdale JT;
 XX
 XX WPI; 2001-316332/33.
 DR
 DR N-PSDB; AAD06206.
 XX
 XX High throughput method for screening for biological agents against fatty acid biosynthesis comprises contacting a bacterial enzymatic pathway with enzymes e.g. malonyl-CoA:ACP transacylase -
 PT
 PT
 PT
 PS Claim 1; Page 10; 94pp; English.
 XX

The present invention relates to a high throughput method for screening biological agents affecting fatty acid biosynthesis, comprises contacting a bacterial enzymatic pathway with enzymes. The method is used for screening biological agents affecting fatty acid biosynthesis. Agonists and antagonists of fab (fatty acid biosynthesis) are used to inhibit, prevent or treat diseases such as infections of the upper respiratory tract (e.g. otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g. empyema, lung abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g. secretory diarrhoea, splenic abscess, retroperitoneal abscess), central

Sequence 246 AA;

Query Match 24.9%; Score 348; DB 22; Length 246;

Best Local Similarity 32.6%: Pred. No. 1.1e-26:
24.9%, SCORE 346; DB 22;
query match

Sequence	Similarity	Seq. ID	Accession	Length	Score	E-value	Identical	Mismatches	Indels	Gaps
1	79	Conservative	57	Mismatches	102				Indels	4
2	79	Conservative	57	Mismatches	102				Indels	4
3	79	Conservative	57	Mismatches	102				Indels	4
4	79	Conservative	57	Mismatches	102				Indels	4
5	79	Conservative	57	Mismatches	102				Indels	4
6	79	Conservative	57	Mismatches	102				Indels	4
7	79	Conservative	57	Mismatches	102				Indels	4
8	79	Conservative	57	Mismatches	102				Indels	4
9	79	Conservative	57	Mismatches	102				Indels	4
10	79	Conservative	57	Mismatches	102				Indels	4
11	79	Conservative	57	Mismatches	102				Indels	4
12	79	Conservative	57	Mismatches	102				Indels	4
13	79	Conservative	57	Mismatches	102				Indels	4
14	79	Conservative	57	Mismatches	102				Indels	4
15	79	Conservative	57	Mismatches	102				Indels	4
16	79	Conservative	57	Mismatches	102				Indels	4
17	79	Conservative	57	Mismatches	102				Indels	4
18	79	Conservative	57	Mismatches	102				Indels	4
19	79	Conservative	57	Mismatches	102				Indels	4
20	79	Conservative	57	Mismatches	102				Indels	4
21	79	Conservative	57	Mismatches	102				Indels	4
22	79	Conservative	57	Mismatches	102				Indels	4
23	79	Conservative	57	Mismatches	102				Indels	4
24	79	Conservative	57	Mismatches	102				Indels	4
25	79	Conservative	57	Mismatches	102				Indels	4
26	79	Conservative	57	Mismatches	102				Indels	4
27	79	Conservative	57	Mismatches	102				Indels	4
28	79	Conservative	57	Mismatches	102				Indels	4
29	79	Conservative	57	Mismatches	102				Indels	4
30	79	Conservative	57	Mismatches	102				Indels	4
31	79	Conservative	57	Mismatches	102				Indels	4
32	79	Conservative	57	Mismatches	102				Indels	4
33	79	Conservative	57	Mismatches	102				Indels	4
34	79	Conservative	57	Mismatches	102				Indels	4
35	79	Conservative	57	Mismatches	102				Indels	4
36	79	Conservative	57	Mismatches	102				Indels	4
37	79	Conservative	57	Mismatches	102				Indels	4
38	79	Conservative	57	Mismatches	102				Indels	4
39	79	Conservative	57	Mismatches	102				Indels	4
40	79	Conservative	57	Mismatches	102				Indels	4
41	79	Conservative	57	Mismatches	102				Indels	4
42	79	Conservative	57	Mismatches	102				Indels	4
43	79	Conservative	57	Mismatches	102				Indels	4
44	79	Conservative	57	Mismatches	102				Indels	4
45	79	Conservative	57	Mismatches	102				Indels	4
46	79	Conservative	57	Mismatches	102				Indels	4
47	79	Conservative	57	Mismatches						

55 KVAELVIA SIDDIGI GATAKRAE² GAHV VS - SRKQUNVDQAVATLQGE

3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 10

[illegible]

1000

100

100

184 pgtlvsmtdal--sdelkeqmltriplarfgqtdiantvaflasdkaak

212 GG 213

242 gg 243

Search completed: January 22, 2002, 15:19:05
Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993-2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:17:30 ; Search time 12.54 Seconds
(without alignments)
498.877 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHKAGLLGICARAWNSVRMA.....DASYITGETVVVGGTPSRL 278

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/2/1aa/PCT05_COMB.pap.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	348	24.9	246	3	US-09-238-481-2
2	335.5	24.0	313	4	US-09-413-814-9
3	328.5	23.5	345	3	US-08-793-035-9
4	328.5	23.5	315	3	US-08-793-035-10
5	324.5	23.2	333	1	US-08-440-856A-4
6	324	23.2	262	4	US-09-363-189B-6
7	321	22.9	247	1	US-08-241-766-13
8	313	22.4	244	1	US-08-375-962B-13
9	313	22.4	244	2	US-08-562-114B-13
10	313	22.4	244	4	US-08-729-594A-13
11	312.5	22.3	256	1	US-08-594-808B-7
12	306.5	21.9	337	1	US-08-440-856A-3
13	299.5	21.4	303	3	US-09-002-298-1
14	296.5	21.2	257	4	US-09-287-097-2
15	295	21.1	244	1	US-08-762-129-1
16	295	21.1	244	2	US-09-090-567-2
17	295	21.0	271	2	US-07-637-865-2
18	294	21.0	244	1	US-08-762-129-3
19	279	19.9	246	6	5229279-7
20	272.5	19.5	255	4	US-08-815-225-4
21	270	19.3	251	3	US-08-822-322-9
22	270	19.3	251	4	US-09-466-109-9
23	270	19.3	252	3	US-08-822-322-8
24	270	19.3	252	4	US-09-466-109-8
25	266	19.0	244	1	US-08-762-129-4
26	258.5	18.5	283	4	US-09-367-012-1
27	253	18.1	273	6	5512669-4

28	248	17.7	359	1	US-08-440-856A-8	Sequence 8, Appli
29	246.5	17.6	248	4	US-09-385-028-11	Sequence 11, Appl
30	241	17.2	335	3	US-09-002-298-7	Sequence 7, Appli
31	237	16.9	295	3	US-09-002-298-5	Sequence 5, Appli
32	230.5	16.5	292	3	US-09-109-205-2	Sequence 2, Appli
33	226	16.2	335	3	US-09-002-298-6	Sequence 6, Appli
34	224	16.0	263	6	5229279-4	Patent No. 5229279
35	221	15.8	335	3	US-09-109-205-19	Sequence 19, Appl
36	217.5	15.5	231	2	US-08-336-198C-7	Sequence 7, Appli
37	217.5	15.5	295	4	US-09-026-482B-2	Sequence 2, Appli
38	197.5	14.1	327	1	US-08-375-962B-12	Sequence 12, Appl
39	197.5	14.1	327	2	US-08-562-114B-12	Sequence 12, Appl
40	197.5	14.1	327	4	US-08-729-594A-12	Sequence 12, Appl
41	191.5	13.7	261	4	US-08-815-225-2	Sequence 2, Appli
42	191.5	13.7	261	4	US-08-815-225-3	Sequence 3, Appli
43	183.5	13.1	388	4	US-08-980-832-41	Sequence 41, Appl
44	173	12.4	325	3	US-08-581-148C-4	Sequence 4, Appli
45	170.5	12.2	309	3	US-09-109-205-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-238-481-2
; Sequence 2, Application US/09238481
; Patent No. 6110704
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: Fabg
; FILE REFERENCE: GM10192
; CURRENT APPLICATION NUMBER: US/09/238,481
; CURRENT FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-238-481-2

Query Match	24.9%	Score 348;	DB 3;	Length 246;
Best Local Similarity	32.6%	Pred. No. 3.8e-30;		
Matches	79;	Conservative	57;	Mismatches 102;
		Indels	4;	Gaps 3;
Qy	33	KVALVTASTDGGFAIARRLAODGAHVVS-SRKOQNVDOAVATLQEGLSVTGTVCHVG	91	
Db	5	KSALVTGASRGIGRSIALQLAEEGYNVAVNAGSKEKAEVVEEIKAKCVDSFAIQANVA	64	
Qy	92	KAEDRERLVAATKLUHGGIDILVSNRAVNPFFGSIIMDVTEEVWDKTLDINVKAPALMTKA	151	
Db	65	DADVKAMKEVVSQFSLDLVNNAGITR-DNLLMRKKEQEMDDVIDNLSKGVFNICOK	123	
Qy	152	VVPEMEKGGSVLVSSTAAPSPGSPYNSVNTALLGLTKTLAIELAPNRVNCIA	211	
Db	124	ATPQMLRQSGAIINLSVVGAVGNPGQANYATKAGVIGLTKSAARELASRGITVNAVA	183	
Qy	212	PCLIKTSFRLMMDKEESMKETLRRRLRCEPCAGINVSFLCSEDASYITGETVVVG	271	
Db	184	PGFIVSDMTDAL--SDELKEQMLTOIPLARFGQDDTIANTVAFLASDKAKYITGOTIHVN	241	
Qy	272	GG 273		
Db	242	GG 243		

RESULT 2
US-09-413-814-9
; Sequence 9, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:

APPLICANT: Gesellschaft für Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bioecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 313
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-9

Query Match 24.0%; Score 335.5; DB 4; Length 313;
Best Local Similarity 33.6%; Pred. No. 1.3e-28;
Matches 83; Conservative 50; Mismatches 105; Indels 9; Gaps 4;

QY 30 LANKVALVASTDGTGFATRLAQDGAHVYSSRRKQNVQAVATLOGEGLSVGTGVC 88
DB 50 LAGRALVATGSSRGICKATLRLAEOGADVAVNHSNKDAAEQTAETRALGRRTMVVQA 109

QY 89 HVGKAEDRLRLVATVKLHGGIDILVSNAAVNPFFGSDVTEEVWDTLDINVKAPALM 148
DB 110 DVTRNAAELFSSVEALGPDIDHVNNGV-DFFPKPLAAMTDDMRNVMSNLSVHYL 168

QY 149 KAVVPEMEKRGSGSVIVYSSIAAFSPGFSFYNNYSKTLGLTKTLAELAPRNR 206
DB 169 CRAAVARMQRKSGRIINIGLSPFAIRCAPNVAYSTAKTGVLTSLATEEAPHGIL 228

QY 207 VNCAPGLIKTSFSEMLWMDKESMKETLRIRLGPEDCAGIVSLFCSDESAYITGE 266
DB 229 VNCVSPGLIDNGY-----LPPAQKEWMRRVPMGRGLGRASEVADAVAFIADRASYVSGA 283

QY 267 TVVGGG 273
DB 284 NIAVAGG 290

RESULT 3
US-09-793-035-9
Sequence 9, Application US/08/793035
Patent No. 6011201
GENERAL INFORMATION:
APPLICANT: Slabas, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Dianne
APPLICANT: Elborough, Keiran
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-ketoacyl ACP Reductase Genes From
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,035
FILING DATE: 28-JUL-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9414622.2
FILING DATE: 20-JUL-1994
APPLICATION NUMBER: GB95/01678
FILING DATE: 17-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kammerer, Patricia A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: MOBT:132
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713.787.1400
TELEFAX: 713.787.1440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-793-035-9

Query Match 23.5%; Score 328.5; DB 3; Length 315;
Best Local Similarity 33.1%; Pred. No. 7.9e-28;
Matches 81; Conservative 53; Mismatches 100; Indels 11; Gaps 5;

QY 34 VALVYASTDGTGFATRLAQDGAHVYV-----SSRQNVQVDOAVATLOGEGLSVGTGVC 89
DB 74 VVVVTGASRGICKATLRLAEOGADVAVNHSNKDAAEQTAETRALGRRTMVVQA 130

QY 90 VGKAEDRLRLVATVKLHGGIDILVSNAAVNPFFGSDVTEEVWDTLDINVKAPALM 149
DB 131 VSKEADVAMKKTADANGTIDVVVNNAGITR-DTLLIRMKKSQWDEVIDLNTGVFLCT 189

QY 150 KAVVPEMEKRGSGSVIVYSSIAAFSPGFSFYNNYSKTLGLTKTLAELAPRNRVNC 209
DB 190 QAAATKIMMKRRGRIINIASVVGGLIGNTQOANYAAKAGVIGFSKTAAREGASRNINNV 249

QY 210 LAPGLIKTSFSEMLWMDKESMKETLRIRLGPEDCAGIVSLFCSDESAYITGETV 268
DB 250 VCPGFIASDMTAKLGEDMEKK--ILGTIPLGRYGOPEDVAGLVEFLALSPAASYITGOAF 307

QY 269 VVGGG 273
DB 308 TIDGG 312

RESULT 4
US-08-793-035-10
Sequence 10, Application US/08/793035
Patent No. 6011201
GENERAL INFORMATION:
APPLICANT: Slabas, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Dianne
APPLICANT: Elborough, Keiran
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-ketoacyl ACP Reductase Genes From
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX


```

ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440.856A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 05463-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1517
TELEFAX: (202) 887-0763
TELEX: 706141
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-440-856A-4

```

Query Match 23.2%; Score 324.5; DB 1; Length 333;
Best Local Similarity 32.9%; Pred. No. 2.4e-27;
Matches 102; Conservative 37; Mismatches 108;
Indels 63; Gaps 11;

13	Qy	ANNSVMASSGTRRDP	LANKVALV	YASTD	TG	FAIR	LAODGA	HVVVSS	SRKQNVQA	72
		:	:	:	:	:	:	:	:	
37	Db	AWDSPNGAPT	MPKR	---	LEGK	VAIT	TGGARG	ICEA	IVRLFV	HGAKVIA
		:	:	:	:	:	:	:	:	89
73	Qy	VATLOGE	GLSV	---	TGTV	---	CHVGK	ADRR	RLVAT	KVLHGG
		:	:	:	:	:	:	:	:	124
90	Db	A	---	GEAL	AAAL	PHVGE	VRCD	SV	VEDR	VAER
		:	:	:	:	:	:	:	:	142
125	Qy	---	SIMDV	TEEV	WDK	TD	IN	VKAL	PMTK	AVV
		:	:	:	:	:	:	:	:	179
143	Db	TRAA	SILSF	DAGE	FD	RV	LR	VAL	GAAL	GK
		:	:	:	:	:	:	:	:	202
180	Qy	SPYNSK	TALL	GLIT	KT	LA	TEL	AP	RI	RV
		:	:	:	:	:	:	:	:	225
203	Db	HAYT	ASKH	AI	VGLT	KN	ACE	EL	AG	IR
		:	:	:	:	:	:	:	:	262
226	Qy	---	DKEE	S	MKE	TLR	---	IR	RG	EP
		:	:	:	:	:	:	:	:	268
263	Db	IDLD	IAVPSDQ	EV	KE	ME	VVR	GL	AT	LG
		:	:	:	:	:	:	:	:	318
269	Qy	VVGGG	---	TPSR	277					
		:	:	:	:					
319	Db	VVDG	GV	TTSR	328					
		:	:	:	:					

```

RESULT      6
US-09-363-189B-6
; Sequence 6, Application US/09363189B
; Patent No. 6242228
; GENERAL INFORMATION:
; APPLICANT: SUGIYAMA, MASAKAZU
; APPLICANT: TONOUCHI, NAOTO
; APPLICANT: SUZUKI, SHUNICHI
; APPLICANT: YOKOZEKI, KENZO
; TITLE OF INVENTION:  XYLITOL DEHYDROGENASE OF ACETIC ACID BACTERIA AND GENE THERE
; FILE REFERENCE:      0010-1024-0
; CURRENT APPLICATION NUMBER: US/09/363,189B
; CURRENT FILING DATE:  1999-07-26
; PRIOR APPLICATION NUMBER: JP10-216047
; PRIOR FILING DATE:    1998-07-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 5
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Gluconobacter oxydans
US-09-363-189B-6

Query Match 23.2%; Score 324; DB 4; Length 262;
Best Local Similarity 30.1%; Pred. No. 1.8e-27;
Matches 78; Conservative 55; Mismatches 104; Indels 22; Gaps 3;

QY 33 KVALVTASTDGTGIFATRLAODGAHVHVVSSRKQNVQAVATLOGEGLSVGTGVCHVK 92
DB 8 KVCVLTGAGNIGLATALALAEETGTAIALDMMNREALEKAEASVEKGEARSVCVDTVS 67
QY 93 AEDRELAVATAVKLHGIDILVSNAAVNPFFGSIIMDVTEEVMDKTLDINVKAPALMTKAV 152
DB 68 EEAIVGTVDVVRDGGKIDFLFNAGYOGAFAPVDYFSDDFARVLTINVTGAFHLKAV 127
QY 153 VPEMERKGGSVVIVSSIAAFSPGFPYVSKTALGLTKTLAIELAPRNIIVNCLAP 212
DB 128 SRQMTQNTGRIVNTASMAVGKPPNMAAYGASKAIIALTETAALDLAPYINIRVNAISP 187
QY 213 GLIKTSFSRLMWDKEKE-----ESMKETLRLRRIGEPEDCAGIVSF 254
DB 188 GTMGPGF---MM-EKQVELQAKVGSYFSTDPKVVAQMGVPMRRYGDINEIPGVVAF 243
QY 255 LCSSEDASYITGETVVVGGG 273
DB 244 LLDGDSFMTGVNLPAGG 262

RESULT 7
US-08-241-766-13
; Sequence 13, Application US/08241766
; Patent No 5685590
; GENERAL INFORMATION:
; APPLICANT: JACOBS, W. R.
; APPLICANT: COLLINS, D. M.
; APPLICANT: BANERJEE, A.
; APPLICANT: GELISLE, G. W.
; APPLICANT: WILSON, T. M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; STREET: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,766
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20003.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 484-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-241-766-13

Query Match 22.9%; Score 321; DB 1; Length 247;
Best Local Similarity 31.9%; Pred. No. 3.6e-27;
Matches 79; Conservative 53; Mismatches 100; Indels 16; Gaps 5;

QY 27 RDLANKVALVYVASTDGTGIFATRLAODGAHVHVVSSRKQNVQAVATLOGEGLSVTGT 86
DB 10 KPFEVSRSVLVTGNGNRIGLATAQLAALADGKHVAVTHRG-----SCAPKGLFCV 58
QY 87 VCHVGKAEDRLERLVATAVKLH--GGIDILVSNAAVNPFFGSIIMDVTEEVMDKTLDINVKAP 145
DB 59 ECDVTDSDAVDR-AFTAVEEHQGPVEVLVSNAGLSA-DAFLMRMTEEFKVKVINANLTGA 116
QY 146 ALMTKAVVPEMERKGGSVVIVSSIAAFSPGFPYVSKTALGLTKTLAIELAPRNI 205
DB 117 FRVAQRASRSRMORNKFGRMIFIGSVSGWIGNQANYAASKAGVIGMARSIARELSKANV 176
QY 206 RVNCLAPGLIKTSFSRLMWDKEKESMKETLRLRRIGEPEDCAGIVSFCSDESASYITG 265
DB 177 TANVAPGVITDMDTRAL--DERIQOQALQFIPAKRVGTPAEVAGVSVFLASEDASYISG 234
QY 266 ETVVVG 273
DB 235 AVIPVDG 242

RESULT 8
US-08-375-962B-13
; Sequence 13, Application US/08375962B
; Patent No 5731195
; GENERAL INFORMATION:
; APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,
; APPLICANT: CHONTER, ERIKSSON, ULF.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecule
; TITLE OF INVENTION: Which Codes for A 32 kDa Protein Having 11-CIS Retinol
; TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With P63
; TITLE OF INVENTION: a Portion of a Retinol Binding Protein Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect (ASCII standard)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,962B
; FILING DATE: 20-January-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,418
; FILING DATE: 6-October-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
;
; FEATURE:
;   NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
;   NAME/KEY:
US-08-562-114B-13

```

Query Match	22.4%;	Score 313;	DB 2;	Length 244;	
Best Local Similarity	32.8%;	Pred. No. 2.6e-26;			
Matches	81;	Conservative 53;	Mismatches 95;	Indels 18;	Gaps
QY	33	KVALVTASTDGTGGAFAARRLAODGAHVHVSSRKQNVDOAVAT---	LOGEGLSVTGTVCH	89	
Db	6	KIALVTGASRGITGRAIETLAARGCK-VIGTATSENGAQAI	SDYLGA	61	
QY	90	VGKAEDRELRIVATAVKLH---	GGIDILVNSAAVNPFFGSDIMDVT	146	
Db	62	-----DPASIESVLEKIRAEFG	DEV	115	
QY	147	LMTKAVVPMEKRGGSVVIVSSIAAFSPGSPVNVSKTALLGLT	KTALTELAPNRIR	206	
Db	116	RLSKAVPAMMKRHRGRIITIGSVGTWNGCQANYAAKAGLIG	FSKSLAREVASRGIT	175	
QY	207	VNCLAPFLIKTFSRMLMWDKESMSKTELRIIRIGEPEDCAGIV	SVFLCSDASVIITGE	266	
Db	176	VNVVAFSGFIETDMTRAL--SDQ	RAGILAQVPA	233	
QY	267	TVVVG	GG	273	
Db	234	TLHVNGG		240	

RESULT 10
US-08-729-594A-13
: Sequence 13, Application US/08729594A
: Patent No. 6280997
: GENERAL INFORMATION:
: APPLICANT: Eriksson, Ulif; Simon, Andras; Romert, Anna
: TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH
: TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGEN
: TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A
: TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felfe & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/729,594A
: FILING DATE: 11-October-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/562,114
: FILING DATE: 22-No. 6280997ember-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/375,962
: FILING DATE: 20-January-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/258,418
: FILING DATE: 10-June-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanson, No. 6280997man D.
: REGISTRATION NUMBER: 30,946
: REFERENCE/DOCKET NUMBER: LUD 5372.3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
NAME/KEY:
US-08-729-594A-13

Query Match 22.48; Score 313; DB 4; Length 244;
Best Local Similarity 32.88; Pred. No. 2.6e-26;
Matches 81; Conservative 53; Mismatches 95; Indels 18; Gaps 6;
QY 33 KVALVTASTDGIQFAIRRLAODGAHVVSRRKQONVDQAVAT---LQEGSLVTGTVCH 89
DB 6 KIALVTGASRGIGRAIAETLAARGK-VIGTATSENGAQALSDYLGANGKGLMLNVT--- 61
QY 90 VGKAEDRERLVAATKLVH---GGIDILVSNAAVNPFFGSIIMDVTEEVMDKTLIDINVKAPA 146
DB 62 -----DPASIESVLEKIRAECEVDILVNNAGITR-DNLMRMKDEWNDIETNLSVF 115
QY 147 LMTKAVVPEMEKRGSGSVVIVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNR 206
DB 116 RLSKAVVRAMMKRHRITIGTSVVTMGNGQYAAAKAGLIGFSKSLAREVASRGIT 175
QY 207 VNCIAPGLIKTSFMRMLMDKEESMKETLIRRLGEPEDCAGIVFLCSEDASYITGE 266
DB 176 VNVVAPGFIEDMTAL--SDDQAGILAQVPAGRLGGAQEIANAVAFSLASDEAAAYITGE 233
QY 267 TVVVGSG 273
DB 234 TLVNGG 240

RESULT 11
US-08-594-808B-7
Sequence 7, Application US/08594808B
Patent No. 5804423
GENERAL INFORMATION:
APPLICANT: Klasen, Ralf
APPLICANT: Bringer-Meyer, Stephanie
APPLICANT: Sahm, Hernand
APPLICANT: Hollenberg, Cornelles P
TITLE OF INVENTION: MICROBIOLOGICAL METHOD OF MAKING
TITLE OF INVENTION: 5-KETOGLUCONATE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Firm of Karl F. Ross, PC
STREET: 5676 Riverdale Ave.
CITY: Bronx
STATE: New York
COUNTRY: USA
ZIP: 10471-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,808B
FILING DATE: 07-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Jonathan E
REGISTRATION NUMBER: 26,963
REFERENCE/DOCKET NUMBER: 19893
TELECOMMUNICATION INFORMATION:
TELEPHONE: (718) 884-6600
TELEFAX: 718/601-1099

TELEX: 620428
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-594-808B-7
Query Match 22.38; Score 312.5; DB 1; Length 256;
Best Local Similarity 31.28; Pred. No. 3.2e-26;
Matches 77; Conservative 53; Mismatches 110; Indels 7; Gaps 3;
QY 30 LANKVALVTASTDGIQFAIRRLAODGAHVVSRRKQONVDQAVATLQEGSLVTGTVCH 89
DB 9 LSGARALVTGASRGIGITLAKGLARYGAEVVLNGRNAESLDSAQSGFEAGKGLASTAVFD 68
QY 90 VGKAEDRERL---VATAVKLHGGIDILVSNAAVNPFFGSIIMDVTEEVMDKTLIDINVKAPA 146
DB 69 V---TDQDAVIDGVAATERDMPIDILINNAGIQR-RAPLEEFERKDDWDLMTSTNVNAV 124
QY 147 LMTKAVVPEMEKRGSGSVVIVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNR 206
DB 125 FVGQAVARHMPRGRKIVNICSVQSELARPGIAPYTATKGAVKNTKGMATDWRHGLQ 184
QY 207 VNCIAPGLIKTSFMRMLMDKEESMKETLIRRLGEPEDCAGIVFLCSEDASYITGE 266
DB 185 INGLAPGYFATEMTERLVADEEFTDNLCKRTPAGRWQGVQVELVGAAVFLSRSSFFVNGQ 244
QY 267 TVVVGSG 273
DB 245 VLMVDGG 251

RESULT 12
US-08-440-856A-3
Sequence 3, Application US/08440856A
Patent No. 5750873
GENERAL INFORMATION:
APPLICANT: DELLAPORTA, STEPHEN L.
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVE. N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,856A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 05463-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1517
TELEFAX: (202) 887-0763
TELEX: 706141
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
STRANDEDNESS: single

QY 148 MTKAVPEMEKRGSGVIVVSSIAAFSPGSPYNVSKTALLGLTKTLAIELAPRNIRV 207
Db 118 TPAALPGRAKGRVIRVIAHAGLITSPYKSAVAAHGVRGFTKVTALGTAGKGITC 177
QY 208 NCLAPGLIKTSF-----SRMLMDKEESMKETL----RIRRLGEPEDCAGIVSFCLCS 257
Db 178 NAICPGYVLTPLVEAQIPDOMKAHMDRETIVREYMLDRQPSRQFATTGQIGGTVVFLCS 237
QY 258 EDASYITGETVVVGGG 273
Db 238 GAADQITGTTISVDGG 253

RESULT 15
US-08-762-129-1
; Sequence 1, Application US/08762129
; Patent No. 5756299
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN CARBONYL REDUCTASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08762,129
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0171 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-08-762-129-1

Query Match 21.1%; Score 295; DB 1; Length 244;
Best Local Similarity 32.7%; Pred. No. 2.5e-24;
Matches 81; Conservative 44; Mismatches 107; Indels 16; Gaps 5;
QY 30 LANKVALVTASTDGIGFAIARRLAQDGAHVYVSSRKQOONDQAVATLQGBGLSVGTGVCH 89
Db 5 LAGRRVLVTGAGRGIGRGVTQALHATGARVAVVSRQTQADLDSLVRCPG----IEPVCDV 60
QY 90 VGRAEDRERLVATVKLHGIDILVNAAV---NPFEGSIMDYETEVDKTLINVKAPA 146
Db 61 LGDWEATERALGSV---GPVDLLVNNAAVALLQPF---LEVTKEAFDRSFEVNLRAVI 112

QY 147 LMTKAVVPEMEKRG--GGSVIVVSSIAAFSPGSPYNVSKTALLGLTKTLAIELAPRNIRV 205
Db 113 QVSOIVARGLLIARGVPGAIVNVSSQCSQRAVTNHSVYCTKGALOMLTKVMALELGPHKI 172
QY 206 RYVCLAPGLIKTSFSPSRMLMDKEESMKETLIRRLGEPEDCAGIVSFCLSEDASYITG 265
Db 173 RYNAVNPVTVMVMTSMGOATWSDPHKAKTMXNRIPXGKFAEVEHVVNAILFLLSDRSGMTTG 232
QY 266 ETVVVVG 273
Db 233 STLPVEGG 240

Search completed: January 22, 2002, 15:19:24
Job time: 114 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002 15:18:05 ; Search time 16.46 Seconds
(without alignments)
1286.545 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHKAGLLGLCARANNVRMADASYITGVVVGTPSPSL 278
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_68
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	824.5	58.9	280	2 S66665	nuclear protein He
2	636	45.5	257	2 T32002	hypothetical prote
3	614	43.9	260	2 S2476	hypothetical prote
4	588.5	42.1	254	2 E5069	hypothetical prote
5	459.5	32.8	260	2 E70881	probable fabG5 pro
6	409	29.2	255	2 D83416	probable short-cha
7	399	28.5	246	2 A69621	3-oxoacyl-(acyl)-ca
8	396.5	28.3	246	2 H72219	3-oxoacyl-(acyl)-ca
9	388	27.7	248	2 C83961	3-oxoacyl-(acyl)-ca
10	385	27.5	248	2 H70447	3-oxoacyl-(acyl)-ca
11	381.5	27.3	282	2 C71204	probable dehydroge
12	381	27.2	255	2 H84288	3-oxoacyl-(acyl)-ca
13	374	26.7	243	2 F86721	hypothetical prote
14	373.5	26.7	251	2 C75217	probable short-cha
15	371	26.5	253	2 C75217	probable short-cha
16	369.5	26.4	285	2 D83767	glucose 1-dehydrog
17	369.5	26.4	285	2 D85885	probable oxidoredu
18	368.5	26.3	285	1 A65017	probable 3-oxoacyl
19	368.5	26.3	320	2 S22450	3-oxoacyl-(acyl)-ca
20	366	26.2	261	2 S35196	hypothetical prote
21	365.5	26.1	248	2 F69868	glucose 1-dehydrog
22	364.5	26.1	260	2 B48674	troponine reductas
23	364	26.0	254	2 G75333	3-oxoacyl-acyl car
24	360.5	25.8	261	2 S54815	urdamycin polyket
25	360	25.7	272	1 S05397	granaticin polyket
26	358.5	25.6	256	2 A70637	hypothetical prote
27	357	25.5	248	2 F81971	probable 3-oxoacyl
28	357	25.5	248	2 E81026	3-oxoacyl-(acyl)-ca
29	357	25.5	273	1 A48674	troponine reductas

30	354	25.3	241	2 F71636	3-oxoacyl reductas
31	354	25.3	278	2 T10877	y4IA protein - Rhi
32	353	25.2	262	2 B84695	probable troponine
33	349	24.9	256	2 E72427	oxidoreductase, sh
34	349	24.9	258	2 G69755	glucose 1-dehydrog
35	347.5	24.8	253	2 E83207	probable short-cha
36	345.5	24.7	288	2 B84693	probable troponine
37	345	24.7	246	1 DEKCN	acetoadetyl-CoA re
38	345	24.7	248	2 D70707	hypothetical prote
39	344.5	24.6	255	2 G82644	2,5-dichloro-2,5-c
40	343	24.5	248	2 F82128	3-oxoacyl-(acyl)-ca
41	342	24.4	247	2 S77280	3-oxoacyl-(acyl)-ca
42	341.5	24.4	246	2 H84136	3-oxoacyl-(acyl)-ca
43	341	24.4	248	2 E81695	3-oxoacyl-(acyl)-ca
44	341	24.4	268	2 E84695	probable troponine
45	341	24.4	281	2 F69400	2-deoxy-D-gluconat

ALIGNMENTS

RESULT 1
S66665
nuclear protein Hep27 - human
N:Alternate names: protein D
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S66665; S66614; S14094; S78122
R:Gabrielli, F.; Donadel, G.; Bensi, G.; Heguy, A.; Melll, M.
A:Title: A nuclear protein, synthesized in growth-arrested human hepatoblastoma cells
A:Reference number: S66614; MUID:96035881
A:Accession: S66665
A:Molecule type: mRNA
A:Residues: 1-280 <GAB>
A:Cross-references: EMBL:U31875; NID:g1079565; PIDN:AAA82048.1; PID:g1079566
A:Note: Met-23 is the probable initiator
A:Accession: S66614
A:Molecule type: protein
A:Residues: 24-41 <GAW>
R:Donadel, G.; Garzelli, C.; Frank, R.; Gabrielli, F.
Eur. J. Biochem. 195, 723-729, 1991
A:Title: Identification of a novel nuclear protein synthesized in growth-arrested human
A:Reference number: S14094; MUID:91153312
A:Accession: S14094
A:Molecule type: protein
A:Residues: 80-83, 'L', '85, 87-88, 'A', '90-91; 141-146; 148-162; 198-205 <DON>
A:Accession: S78122
A:Molecule type: protein
A:Residues: 'V', '81-84, 'G', '86-88 <DOW>
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: nucleus
F:37-216/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match	58.9%	Score	824.5	DB 2	Length	280			
Best Local Similarity	61.8%	Pred. No.	3.4e-58						
Matches	162	Conservative	47	Mismatches	51	Indels	3	Gaps	2
QY	10	CARAWNSVRMASSGMRDRDPLANKVALVTASTDGGFAIARRLAQDGAHVWVSSRKQONV	69						
Db	16	CARL--SVRMSSTGDRKGVLANRVAVVTGSGIGFAIARRLAQDGAHVWVSSRKQONV	73						
QY	70	DQAVATLQGEGLSVGTGTVCHVGKAEERLRVATKVLHGDDILVSNAAVNPFGSMDV	129						
Db	74	DRAMAKLQGEGLSVAGIVCHVGKAEERLRVATKVLHGCGVDFVLCAGVNPVVGSLGT	133						
QY	130	TEEVWDTLDINKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPGSGFPYNNVKTAL	189						
Db	134	SEQIWDRIKLVNVKSPALLSQQLPYMENR-RGAVILVSSIAAYNPVVALGVNVSKTAL	192						
QY	190	LGLTKTILAEIAPNIRVNCIAPGLIKTSFRLMMDKEESMKETLRIRRLGEPDCA	249						

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-246 <KUN>
A:Cross-references: GB:2799112; GB:AL009126; NID:g2633902; PIDN:CAB13464.1; PID:g26339
A:Experimental source: strain 168
R:Oguro, A.; Kakeshita, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996
A:Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit,
A:Reference number: JC4819; MUID:96257247
A:Accession: PC4176
A:Molecule type: DNA
A:Residues: 230-246 <OGU>
A:Cross-references: DDBJ:D64116; NID:gl389548; PIDN:BAA10974.1; PID:gl1237012
R:Cronan, J. E.; Morbidoni, H. R.; de Mendoza, D.

J. Bacteriol. 178, 4794-4800, 1996
 A:Title: Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosyn-
 A:Reference number: 223107; MUID:96326321
 A:Accession: T46633
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22, 'A', 24-246 <CRO>
 A:Cross-references: EMBL:059433; NID:gl502418; PIDN:AAC44307.1; PID:gl502421
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: fabG; srb
 A:Map position: 135-145 degrees
 C:Function:
 A:Description: EC 1.1.1.100 [validated; MUID:96326321]
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NADP; oxidoreductase
 F:5-185/Domain: short-chain alcohol dehydrogenase homology <SADH>
 F:154/Active site: Tyr #status predicted

Query Match 28.5%; Score 399; DB 2; Length 246;
 Best Local Similarity 37.6%; Pred. NO. 1.8e-24;
 Matches 92; Conservative 50; Mismatches 99; Indels 4; Gaps 3;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGEGLSVTGTVC 88
 Db 2 LNDKTAIVTGASRGIGKSTALDIAKSGANVYVNSNEAKANEVDEIKSMGRKATAVKA 61
 QY 89 HVGKAEDRRLVATAVKLHGIDILVSNAAVNPFGSINDVTEVVDKTLIDINVKAPALM 148
 Db 62 DVSNPEDVQNMIKETLSVSTIDILVNNAGITR-DNLIMRKEDEWDDVINILKGVFNC 120
 QY 149 TRAVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRNIRVN 208
 Db 121 TRAVTRQMKQRSGRIINNVSSVGVSGNPGQANYAAKAGVIGLTKSSAKELASRNITVN 180
 QY 209 CLAPGLIKTSFMRMLWMDKEESMKETLIRRLGPEPCAGIVSFLCSEDASYITGETV 268
 Db 181 ATAPGISTDMTDKL--AKVDQDEMLKQPLARFGPESDVSSVYVTFPLASEGARYMTGTL 238
 QY 269 VVGGG 273
 Db 239 HIDGG 243

RESULT 8
 H72219
 3-oxoacyl-(acyl carrier protein) reductase - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: H72219
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316
 A:Accession: H72219
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <ARN>
 A:Cross-references: GB:AE001811; GB:AE000512; NID:g4982291; PIDN:AAD36790.1; PID:g498230
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1724
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 F:6-185/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 28.3%; Score 396.5; DB 2; Length 246;
 Best Local Similarity 36.5%; Pred. NO. 2.8e-24;
 Matches 89; Conservative 54; Mismatches 198; Indels 3; Gaps 3;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGEGLSVTGTVC 89
 Db 3 LEGKVLITGAASGIGKATLLFAQEGATVIAGDISKENLDSLVKEAGLPGKVPVVLN 62
 QY 90 VGKAEDRRLVATAVKLHGIDILVSNAAVNPFGSINDVTEVVDKTLIDINVKAPALMT 149
 Db 63 VDRDOIKEVKEVKVQYGRIDLVLNNAGITR-DALLVRMKEEDMDAVINVLKGVNFVT 121
 QY 150 KAVVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRNIRVNC 209
 Db 122 QMVVPYMIKORNGSIVNVSVVGIYNGPQTNAAKAGVIGMTKTWAKELAGNIRVNA 181
 QY 210 LAPGLIKTSFMRMLWMDKEESMKETLIRRLGPEPCAGIVSFLCSEDASYITGETV 269
 Db 182 VAPGFIETPMTKEL-PEKARETALS-R-IPLGRFGKPEEVAQVILFLASDESSVYTGQVIG 239
 QY 270 VGGG 273
 Db 240 IDGG 243

RESULT 9
 C83961
 3-oxoacyl-(acyl-carrier protein) reductase fabG [imported] - Bacillus halodurans (str
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
 C:Accession: C83961
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: AB3650; MUID:20263314
 A:Accession: C83961
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <STO>
 A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA806210.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: fabG

Query Match 27.7%; Score 388; DB 2; Length 246;
 Best Local Similarity 37.9%; Pred. NO. 1.3e-23;
 Matches 94; Conservative 48; Mismatches 96; Indels 10; Gaps 4;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGEGLSVTGTVC 88
 Db 2 LOGKTAIVTGASRGIGKATAMELARHGANNVYVNSAGNKEKVAEIKELGVEATAIOA 61
 QY 89 HVGKAEDRRLVATAVKLHGIDILVSNAAV--NPFGSINDVTEVVDKTLIDINVKAP 145
 Db 62 DVADSESQVAMKEIDTFGAVDILVNNAGITRDNLF----MRMKEEDMDAVIDNLKGV 117
 QY 146 ALMTKAVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRNI 205
 Db 118 FHCRAVTRPMKQRFGRINNVSSVGVVAGNAGQANYAAKAGVIGLTKTLARELANRNI 177
 QY 206 RVNCLAPGLIKTSFMRMLWMDKEESMKETLIRRLGPEPCAGIVSFLCSEDASYITG 265
 Db 178 TVNAVAPGFIEDMTGEL--PEDVKAQMLGQIPLARLQPEEVAKAVRFLASDASLYLTG 235
 QY 266 ETVVVGGG 273
 Db 236 QTIHNGG 243

RESULT 10
 H70447
 3-oxoacyl-(acyl-carrier-protein) reductase (EC 1.1.1.100) - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
 C:Accession: H70447

C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: F86721
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrlich
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: F86721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-243 <STO>
A:Cross-references: GB:AE005176; NID:g12723693; PIDN:AAK04872.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: fabG1

Query Match 26.7%; Score 374; DB 2; Length 243;
Best Local Similarity 34.4%; Pred. No. 1.7e-22;
Matches 84; Conservative 62; Mismatches 92; Indels 6; Gaps 5;
QY 30 LANKVALVTASTDGGICFAIARLAODGAHVYSSRRQQNVDOAVATLQEGLSVTGTCH 89
Db 3 IKKNVFVTGTRGICAKAIAQAKAGSNLIINGRAIS-EELLAEFTAYGVKAVGISGD 61
QY 90 VGKAEDRELRLVATAVKLHGIDILVSNAAVNPFFGSMIDVTEVWDKTLIDINVKAPALMT 149
Db 62 ISKSEDAKQWAEAIETGLGSDVILVANNAGITR-DGLSLKMSDEDFESVLKINLTGAFNMT 120
QY 150 KAVVPEMEKRGGSVVIVSSIAAFSPGSPFNYSKTKALLGLTKTLAEIAPRNRVNC 209
Db 121 QAVLKPTWTRARGAIIINISVGLMGNAGQANYAASKAGLIGLTKSIAREVAARNVRVA 180
QY 210 LAPGLIKTSFRLMMDKEESMKETLRIIRLGEPCDAGIVSFLCSEDASVITGETV 269
Db 181 VAPGFIESDTEVL-SDKVK-DAMKQIIPKRGFMEETATATQFLAEOE--YMTGQVLT 236
QY 270 VGGG 273
Db 237 IDGG 240

RESULT 14
G72389
Oxidoreductase, short chain dehydrogenase/reductase family - Thermotoga maritima (strain
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72389
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: G72389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-251 <ARN>
A:Cross-references: GB:AE001714; GB:AE000512; NID:g4980823; PIDN:AAD35412.1; PID:g498082
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0325
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.7%; Score 373.5; DB 2; Length 251;
Best Local Similarity 36.5%; Pred. No. 1.9e-22;
Matches 91; Conservative 52; Mismatches 99; Indels 13; Gaps 5;

QY 33 KVALVTASTDGGICFAIARLAODGAHVYSSRRQQNVDOAVATLQEGLSVTGTCH 89
Db 6 KVVLTGAGSGIGKKAAMFAERGAKVAINDISEKGTVELIKSMGGEAAFIQGDV-- 63

QY 90 VGKAEDRELRLVATAVKLHGIDILVSNAAVNPFFGSMIDVTEVWDKTLIDINVKAPALMT 149
Db 64 ---AKDAEQIVKVTETFGRLDILVNNAGIVP-YGNIETSEEDFDKTMVAVNVKGPFLS 119
QY 150 KAVVPEMEKRGGSVVIVSSIAAFSPGSPFNYSKTKALLGLTKTLAEIAPRNRVNC 209
Db 120 KYAVEQMKOGGVIIVNVSSEAGLIGIPRCVYSVSKAALLGLTRSLAVDYDYGIRVA 179
QY 210 LAPGLIKTS--FSRLMMDKEESMKETLRI--RRLGPEPCDAGIVSFLCSEDASVITG 265
Db 180 VCPGTTQSEGLMARVKASPNPELLKKMTRSPMKRLGKEEETAFALLFAACDEAGFMVG 239
QY 266 ETVVVGGGT 274
Db 240 SIINDGGS 248
RESULT 15
C75217
probable short-chain dehydrogenase/reductase PAB2177 - Pyrococcus abyssi (strain Orsa
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75217
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: C75217
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <KAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49186.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2177
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F:9-188/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 26.5%; Score 371; DB 2; Length 263;
Best Local Similarity 36.2%; Pred. No. 3.2e-22;
Matches 92; Conservative 53; Mismatches 97; Indels 12; Gaps 5;

QY 30 LANKVALVTASTDGGICFAIARLAODGAHVYSSRRQQNVDOAVATLQEGLSVTGTVC 88
Db 6 LSGKLAFTTASSKGGIGFVARVLAKAGADVILSRNEENLKAREKIKEESNVVHYIVA 65
QY 89 HVGKAEDRELRLVATAVKLHGIDILVSNAAVNPFFGSMIDVTEVWDKTLIDINVKAPALM 148
Db 66 DLTKMDDLERTV-KELKNIGEPDIFFEFTG-GPKPGYFEMESMEDWEAAVKLLYPAYLV 123
QY 149 TKAVVPEMEKRGGSVVIVSSIAAFSPGSPFNYSKTKALLGLTKTLAEIAPRNRVN 208
Db 124 TKALVPAMERKGFGRITVSTVAIKEIPNIALSNVRIAMAGLVRTLAKELGPKGITVN 183
QY 209 CLAPGLIKTSFRLMMDKEE-----ESMKR---TLRIIRLGEPCDAGIVSFLCSED 259
Db 184 GIMPGLITDRMIQADRAKREGKTVTEALQEQYAKPIPLGRIGPEEIGYLVAFLASDL 243
QY 260 ASYITGETVVVGGG 273
Db 244 GSYINGAMIPVDGG 257

Search completed: January 22, 2002, 15:19:47
Job time: 102 sec

GenCore version: 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:19:25; Search time 9.93 Seconds
(without alignments)
1026.467 Million cell updates/sec

Title: US-09-866-034-2

Perfect score: 1399

Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGCTPSRL 278

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	800.5	57.2	257	1 HE27_HUMAN	Q13268 homo sapien
2	399	28.5	246	1 FABG_BACSU	P51831 bacillus su
3	396.5	28.3	246	1 FABG_THEMA	O9x248 thermotoga
4	385	27.5	248	1 FABG_AQUAE	O67610 aquifex aeo
5	379.5	27.1	260	1 TRN2_HYONI	P50164 hyoscyamus
6	373.5	26.7	231	1 Y325_THEMA	Q9wy90 thermotoga
7	368.5	26.3	320	1 FABG_CUPLA	P28643 cuphea lanc
8	364.5	26.1	260	1 TRN2_DATST	P50163 datura stra
9	362.5	25.9	263	1 UCPA_ECOLI	P37440 escherichia
10	360	25.7	272	1 DHK1_STRVN	P16542 streptomyce
11	357	25.5	273	1 TRN1_DATST	P50162 datura stra
12	354	25.3	241	1 FABG_RICPR	P50941 rickettsia
13	354	25.3	278	1 Y4LA_RHISN	P55541 rhizobium s
14	349	24.9	256	1 Y019_THEMA	Q56318 thermotoga
15	349	24.9	258	1 DHG2_BACSU	P80869 bacillus su
16	345	24.7	246	1 NODG_AZOB	P17611 azospirilli
17	343	24.5	244	1 FABG_VIBCH	Q8kqh7 vibrio chol
18	342	24.4	247	1 FABG_SYNY3	P73574 synechocyst
19	341	24.4	248	1 FABG_CHLMO	Q9pkf7 chlamydia m
20	340	24.3	242	1 FABG_HAEIN	P43713 haemophilus
21	337	24.1	247	1 FABG_CHLTR	P38004 chlamydia t
22	336.5	24.1	256	1 BUDC_KLEPN	Q48436 klebsiella
23	336.5	24.1	261	1 NODG_BACME	P39483 bacillus me
24	336	24.0	245	1 DHG2_RHIS3	P72332 rhizobium s
25	335	23.9	261	1 DHKR_STRCM	P41177 streptomyce
26	334	23.9	267	1 HDHA_CLOSO	P50200 clostridium
27	333.5	23.8	248	1 FABG_CHLPN	O9z8p2 chlamydia p
28	333.5	23.8	255	1 YWFD_BACSU	P39640 bacillus su
29	332	23.7	249	1 DHK2_STRVN	P16543 streptomyce
30	331.5	23.7	261	1 DHG_BACME	P40288 bacillus me
31	331.5	23.7	261	1 DHG_BACSU	P12310 bacillus su
32	330	23.6	247	1 YD50_MYCTU	Q11020 mycobacteri
33	329.5	23.6	261	1 DHG1_BACME	P39482 bacillus me

RESULT 1

HE27_HUMAN STANDARD; PRT; 257 AA.

AC Q13268;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HEP27 PROTEIN (PROTEIN D).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.

RX MEDLINE=96035881; PubMed=7556196;

RA Gabrielli F., Donadel G., Bensi G., Heguy A., Meili M.;

RT "A nuclear protein, synthesized in growth-arrested human

hepatoblastoma cells, is a novel member of the short-chain alcohol

dehydrogenase family.";

RL Eur. J. Biochem. 232:473-477(1995).

RN [2]

RP SEQUENCE OF 57-65; 118-123; 125-139 AND 175-182.

RX MEDLINE=91153312; PubMed=1847869;

RA Donadel G., Garzelli C., Frank R., Gabrielli F.;

RT "Identification of a novel nuclear protein synthesized in growth-

arrested human hepatoblastoma HepG2 cells.";

RL Eur. J. Biochem. 195:723-729(1991).

CC -!- FUNCTION: MAY INHIBIT CELL REPLICATION EITHER BY CATALYSING THE

OXIDATION OF ESTROGEN AND ANDROGEN OR BY CONVERTING CORTISONE IN

CORTISOL.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

(SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U31875; AAA82048.1; ALT_INIT.

DR HSSP; P25529; 1AHH.

DR InterPro; IPR002198; ADH_short.

DR InterPro; IPR002347; Adh_short_C2.

DR Pfam; PF00106; adh_short; 1.

DR Pfam; PF00678; adh_short_C2; 1.

DR PRINTS; PR00081; GDRDH.

DR PROSITE; PS00061; ADH_SHORT; 1.

KW Oxidoreductase; Nuclear protein.

FT INIT_MET 0

FT NP_BIND 17 41

FT ACT_SITE 162 162

FT CONFLICT 57 57

FT CONFLICT 62 62

NAD OR NADP (BY SIMILARITY).

BY SIMILARITY.

L -> V (IN REF. 2).

L -> G (IN REF. 2).

SQ SEQUENCE 257 AA; 27307; 612298E883725CAC CRC64;

Query Match 57.2%; Score 800.5; DB 1; Length 257;
 Best Local Similarity 61.3%; Pred. No. 2.2e-52;
 Matches 155; Conservative 47; Mismatches 50; Indels 1; Gaps 1;

QY 20 ASSGTRRPLANKVALVTASTDGIGFATARRLAQDGAHVVSRRKQNVQDQAVATLQGE 79
 DB 1 STGIDIRKGLNRAVAVVTGSGIGFATARRLAQDGAHVVSRRKQNVQDQAVATLQGE 60

QY 80 GLSVGTGCHVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTL 139
 DB 61 GLSVAGIVCHVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTL 120

QY 140 INVKAPALMTKAVVPMERKGGVSVVSSIAAFSPGFSYVNSKTKALLGLTKTLATE 199
 DB 121 VNVKSPALLSOLLPMENR-RGAVILVSSIAAYNPVVALGVYVNSKTKALLGLTKTLATE 179

QY 200 LAPNIRVNCVAPGLIKTSFSLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSED 259
 DB 180 LAPKDIRVNCVAPGLIKTSFSLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSED 239

QY 260 ASYITGETVVGG 272
 DB 240 ASYVNGENAVAG 252

RESULT 2

FABG_BACSU STANDARD; PRT; 246 AA.

ID FABG_BACSU STANDARD; PRT; 246 AA.

AC P51831.031733;

DT 01-OCT-1996 (Rel. 3.4) Created)

DT 15-DEC-1998 (Rel. 3.7; Last sequence update)

DT 20-AUG-2001 (Rel. 4.0; Last annotation update)

DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
 DE ACYL CARRIER PROTEIN REDUCTASE)

CN FABG.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=96326321; PubMed=8759840;

RA Morbidoni H.R., de Mendoza D., Cronan J.E. Jr.

RT "Bacillus subtilis acyl carrier protein is encoded in a cluster of
 lipid biosynthesis genes."

RL J. Bacteriol. 178:4794-4800(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 1-172 FROM N.A.

RC STRAIN=168;

RX MEDLINE=98195738; PubMed=9534248;

RA Foulger D., Errington J.;

RT "A 28 kbp segment from the spovm region of the Bacillus subtilis 168
 genome."

RL Microbiology 144:801-805(1998).

RN [4]

RP SEQUENCE OF 230-246 FROM N.A.

RC STRAIN=168;

RA Oguro A., Kakeshita H., Nakamura K., Yamane K.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
 NADP(+) -> 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.

CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

CC (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

DR EMBL; U59433; AAC44307.1; -
 DR EMBL; Z99112; CAB13464.1; -
 DR EMBL; Y13937; CAA74250.1; -
 DR EMBL; D64116; BAA10974.1; -
 DR HSSP; Q12634; 1YBV.
 DR Subtilisin; BG11535; fabG.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PRINTS; PR00081; GDRDH.
 DR PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
 DR Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 9 33 NADP (BY SIMILARITY).
 FT ACT_SITE 154 154 BY SIMILARITY.
 FT CONFLICT 23 23 D -> A (IN REF. 1).
 SQ SEQUENCE 246 AA; 26282 MW; C6A391167D3327DC CRC64;

Query Match 28.5%; Score 399; DB 1; Length 246;
 Best Local Similarity 37.6%; Pred. No. 9.5e-23;
 Matches 92; Conservative 50; Mismatches 99; Indels 4; Gaps 3;

QY 30 LANKVALVTASTDGIGFATARRLAQDGAHVVS-SRKQNVQDQAVATLQGEGLSVGTVC 88
 DB 2 LNDKTAIVTASRGIGRSLDLAKSAGVNVYSGNEAKANEVDEIKSMGRKAIVKA 61

QY 89 HVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTLDNVKA PALM 148
 DB 62 DVSNPEDVQNMIKETLSVFSITDILVNNAGITR-DNLIMRKEDWDVNNLKGVCNC 120

QY 149 TKAVVPEMEKGGSVIVSSIAAFSPGFSYVNSKTKALLGLTKTLAELAPNRVNV 208
 DB 121 TKAVTRQMKKORSGRINVSIVSGVSGNPGQANYAAKAGVIGLTKSKAKELASRNITVN 180

QY 209 CLAPGLIKTSFSLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSEDASYITGETV 268
 DB 181 ATAPGFISTDMTDKL--AKDVQDEMLKQIPLARFGEPSDVSVVTFLLASEGARYVTGQTL 238

QY 269 VVGG 273
 DB 239 HIDGG 243

RESULT 3

FABG_THEME

ID FABG_THEME

AC Q9X248;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
 DE ACYL CARRIER PROTEIN REDUCTASE)

GN FABG OR TM1724.

OS Thermotoga maritima.

OC Bacteria; Thermotogales; Thermotoga.

OX NCBI_TaxID=2336;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=99287316; PubMed=10360571;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,

DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TROPINONE REDUCTASE-II (EC 1.1.1.236) (TR-II)
GN TR2.
OS *Hyoscyamus niger* (Henbane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Hyoscyamus.
OX NCBI_TaxID=4079;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cultured root;
RX MEDLINE=94120020; PubMed=8290643;
RA Nakajima K., Hashimoto T., Yamada Y.;
RT "cDNA encoding tropinone reductase-II from *Hyoscyamus niger*."
RL Plant Physiol. 103:1465-1466(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Nakajima K., Hashimoto T.;
RT "Tropinone reductase-II gene of *Hyoscyamus niger*."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE STEREOSPECIFIC REDUCTION OF TROPINONE TO
CC PSEUDOTROPINE.
CC -!- CATALYTIC ACTIVITY: PSEUDOTROPINE + NADP(+) = TROPINONE + NADPH.
CC -!- PATHWAY: BIOSYNTHETIC PATHWAY OF TROPANE ALKALOIDS.
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L20485; AB09776.1;
CC EMBL: AB026545; BA85845.1;
CC HSP: P47227; IBD8.
CC InterPro: IPR002198; ADH_short.
CC InterPro: IPR002347; Adh_short_C2.
CC Pfam: PF00106; adh_short; 1.
CC Pfam: PF00678; adh_short_C2; 1.
CC PRINTS: PRO0081; GDHRDH.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC OXidoreductase; NADP.
CC NP_BIND 13 37 NADP (BY SIMILARITY).
CC FT ACT_SITE 159 159 BY SIMILARITY.
CC SQ SEQUENCE 260 AA; 28437 MW; 6CA7AF85CAA128FC CRC64;

Query Match 27.1%; Score 379.5; DB:1; Length 260;
Best Local Similarity 35.4%; Pred. No. 2.8e-21;
Matches 90; Conservative 49; Mismatches 110; Indels 5; Gaps 3;

QY 24 MTRDPLANKVALVASTDGIIGFAIRLAQDGAHVVSRRKQNVQAVATLOGEGLSV 83
DB 1 MAGRWNEGCTALVTGSGRGIGYIVELANLGASVYTCRSNRKDELCTLRKSGFNV 60

QY 84 TGTVCHVGKAEDERLVATAVK-LHGIDILVSNAAVPPFGSIMDVTVEEWDKTLINV 142
DB 61 EASVCDLSSRSEREEMKTYNHFHGLNVLNAGI-VIYKEAKDTMEDYSHIMSINP-119

QY 143 KAPALMTKAVVPEMEKRGGSVIVSSIAAFSPGSPYNVSKTALLGLTKTIAELAP 202
DB 120 EAAVHLSVAHPFLKASERNVYFISISGASALPYEAVYGATKGAMDQTRCLAFEMAK 179

QY 203 RNRVNCVLAFLGLTKTSRMLWMDKEESMKETL---RIRRGEPEDCAGIVSFLCSED 259
DB 180 DNRVNGVGPVATSVEMTIDPEQKENLDKLDRCALRRMGEPKELAAVAVFLCPA 239

QY 260 ASYITGETVVVGG 273
DB 240 ASYVIGQIIIVDGG 253

RESULT 6

Y325_THEME
ID Y325_THEME STANDARD; PRT; 251 AA.
AC QWYGO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL OXIDOREDUCTASE TM0325 (EC 1.1.1.236).
GN TM0325.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of *Thermotoga maritima*."
RL Nature 399:323-329(1999).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001714; AAD35412.1;
CC TIGR: TM0325.
CC InterPro: IPR002198; ADH_short.
CC InterPro: IPR002347; Adh_short_C2.
CC Pfam: PF00106; adh_short; 1.
CC Pfam: PF00678; adh_short_C2; 1.
CC PRINTS: PRO0080; SDRFAMILY.
CC PRINTS: PRO0081; GDHRDH.
CC PRINTS: PRO1167; INSADHFAMILY.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC Hypothetical protein; Oxidoreductase; Complete proteome.
CC NP_BIND 10 34 NADP (BY SIMILARITY).
CC FT ACT_SITE 152 152 BY SIMILARITY.
CC SQ SEQUENCE 251 AA; 26635 MW; 61C736A0F13564A5 CRC64;

Query Match 26.7%; Score 373.5; DB:1; Length 251;
Best Local Similarity 36.5%; Pred. No. 7.4e-21;
Matches 91; Conservative 52; Mismatches 93; Indels 13; Gaps 5;

QY 33 KVALVTASTDGIGFAIRLAQDGAHV---VSSRKQNVQAVATLOGEGLSVGTGVCH 89
DB 6 KVLVITAGSGIGKKAAMFAERGAKVAINDISEKGETVELIKSMGGGAFFIGDV-- 63

QY 90 VGKAEDRERLVAVAKLHGIGDILVSNAAVPPFGSIMDVTVEEWDKTLINVKAPALMT 149
DB 11 AKAEQIVKKTIVETFGRLDILVNNAGIVP-YGNLEETSEEDFDKTMVAVNKGFLLS 119

QY 150 KAVVPEMEKRGGSVIVSSIAAFSPGSPYNVSKTALLGLTKTIAELAPNRVNC 209
DB 120 KYAVEQMKGGGVIVNVSSSEAGLIGIPRCVYSVSKAALLGLTSLAVDVIDVGIRVA 179

QY 210 LAPGLIKTS--FSRMLWMDKEESMKETLRI--RRIGEPEDCAGIVSFLCSEDASYITG 265
DB 180 VCPGTTQSEGLMARVKASPNPEELLKKTTSRIPMKRLGKEEIAFALFAACDEAGFWTG 239


```
CC EMBL; L20474; AA33282.1;
DR PDB; 2AE1; 18-NOV-98.
DR PDB; 2AE2; 02-FEB-99.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PRINTS; PR00081; GDRDH.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Oxidoreductase; NADP; 3D-structure.
FT NP_BIND 13 37 NADP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 260 AA; 28311 MW; 2DBF4963B2CCA303 CRC64;

Query Match 26.1%; Score 364.5; DB 1; Length 260;
Best Local Similarity 34.6%; Pred. No. 3.6e-20;
Matches 88; Conservative 50; Mismatches 111; Indels 5; Gaps 3;

Oy 24 MTRRDLANKVALVASTDGGIGFAIRLAQDGAHVYSSRKQNVQAVATLOGEGLSV 83
Db 1 MAGRNLEGCTALVTGSGRGIGYVEELASLGASVYTCRNQKELNDCLQWRSGFKV 60

Oy 84 TGTCHVGKAEDEERLVAAT-AVKLHGDDILVSNAAVNPFFGSGIMDVTEYVMDKTLIN 142
Db 61 EASVCDLSRSRQELMTVANHPHGKLNILVNNAGI-VIYKEADYTVEDYSLIMSINF 119

Oy 143 KAPALMTKAVVPEMKRGSGSVYVSSAAFPSPGSPFPYNNVSKTALGLTKTLAIELAP 202
Db 120 EAAVHLSVLAHPFLKASRGNVVTFSSVSGALVPEYAVYGATKGAMPDLTRCLAFENAK 179

Oy 203 RNIRVNCPLAGLTKTSPFRLMMDKEESMKEL--RIRLGPEDCAGIVSFLCSED 259
Db 180 DNRVNGVGPVATSLVEMTIQDPEQENLKLDRCALRRMGPEKELAAWAFVLCPPA 239

Oy 260 ASYITGETVYVGGG 273
Db 240 ASYVTGQIIYVDGG 253

RESULT 9
UCPA_ECOLI STANDARD; PRO: 263 AA.
AC P37440; P76963; P77140;
DT 01-OCT-1994 (Rel. 30, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE OXIDOREDUCTASE UCPA (EC 1.1.1.1)
GN UCPA OR B2426
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett, G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
```

```
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features."
RL DNA Res. 4:91-113(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97385354; PubMed=9241368;
RA Sirko A., Weglenska A., Hryniewicz M.M., Hulanicka D.M.;
RT "Characterization of the Escherichia coli gene encoding a new member
RT of the short-chain dehydrogenase/reductase (SDR) family."
RL Acta Biochim. Pol. 44:153-157(1997).
RN [4]
RP SEQUENCE OF 180-263 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=90264335; PubMed=2189959;
RA Hryniewicz M.M., Sirko A., Palucha A., Boeck A., Hulanicka D.M.;
RT "Sulfate and thiosulfate transport in Escherichia coli K-12:
RT identification of a gene encoding a novel protein involved in
RT thiosulfate binding."
RL J. Bacteriol. 172:3358-3366(1990).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome."
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000330; AAC75479.1; ALT_INIT.
DR EMBL; D90872; BAA16309.1; ALT_INIT.
DR EMBL; D90871; BAA16300.1; -.
DR EMBL; X99908; CAA68181.1; -.
DR EMBL; M32101; -. NOT_ANNOTATED_CDS.
DR HSSP; P14061; IFDV.
DR EcoGene; EGI2133; ucpA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PRINTS; PR00081; GDRDH.
DR PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
KW Oxidoreductase; Complete proteome.
FT CONFLICT 153 153 T -> Q (IN REF. 3).
SQ SEQUENCE 263 AA; 27850 MW; 2E201713357FDF41 CRC64;
```

```
Query Match 25.9%; Score 362.5; DB 1; Length 263;
Best Local Similarity 34.1%; Pred. No. 5.1e-20;
Matches 86; Conservative 53; Mismatches 104; Indels 9; Gaps 4;
```

```
Oy 30 LANKVALVASTDGGIGFAIRLAQDGAHVYSSRKQNVQAVATLOGEGLSVTGTVCH 89
Db 4 LTGXTALITGALQIGEGIAITFAHGHANLILLDISPE-IEKLADELGGRHCTAVVAD 62

Oy 90 VGKAEDEERLVAATVKLHGDDILVSNAAVNPFFGSGIMDVTEYVMDKTLINVKAPALMT 149
Db 63 VRDPASVAAAIKRAKEGRIDILVNNAGVCR-LGSLFLOMSDDDDDFIDINIKGWNVT 121

Oy 150 KAVPEMEKRGSGSVVIVSSIAA-FSPSPGSPYNNVSKTALLGLTKTLAIELAPNIRVN 208
```

Db	76	ADPAQIRAYVAAAVQRYGTVDILVNAG-RSGGGATAEIADELWLVDVITTNLTSPFLMTK	134
Qy	151	AVVPE-----MEKRG-----CGSVVISSIAAFSPCFSPYNSVKTKALLGLTTLAI	198
Db	135	EVLNAGGMLAKKRGIINIASTGQGVHAV-----PYSASKHGVGLTKALGL	184
Qy	199	ELAPRNRVNCIAPGLIKTSFSRM-----LMDKEKBESMKETLRI--RRLGEPEDC	249
Db	185	ELARTGITVNAVCPGFVETPMAERVREHVAGIQVSEETFDITNRPLGRYVETREVA	244
Qy	250	GIYSFLCSEDAASYTGETVVVGGG	273
Db	245	AMVEYLVADDAAAVTAQALNVCVG	268

```

RESULT 11
TRN1_DATST
ID TRN1_DATST STANDARD; PRN; 273 AA.
AC P50162;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TROPINONE REDUCTASE-I (EC 1.1.1.206) (TR-1) (TROPINE DEHYDROGENASE).
GN TR1.
OS Datura stramonium (Jimsonweed) (Common thornapple).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Datura.
OX NCBI_TaxID=4076;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=ROOT;
RX MEDLINE=94022421; PubMed=8415746;
RA Nakajima K., Hashimoto T., Yamada Y.;
RT "Two tropinone reductases with different stereospecificities are
RT short-chain dehydrogenases evolved from a common ancestor.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:9591-9595(1993).
RN [2]
X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=98226735; PubMed=9560196;
RA Nakajima K., Yamashita A., Akama H., Nakatsu T., Kato H.,
RA Hashimoto T., Oda J., Yamada Y.;
RT "Crystal structures of two tropinone reductases: different reaction
RT stereospecificities in the same protein fold.";
RN Proc. Natl. Acad. Sci. U.S.A. 95:4876-4881(1998).
CC -1- FUNCTION: CATALYZES THE STEREOSPECIFIC REDUCTION OF TROPINONE TO
CC TROPINE.
CC -1- CATALYTIC ACTIVITY: TROPINE + NADP(+) = TROPINONE + NADPH.
CC -1- PATHWAY: BIOSYNTHETIC PATHWAY OF TROPANE ALKALOIDS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR	PDB; LAE1; 18-NOV-98.	
DR	InterPro; IPRO02198; ADH_short.	
DR	InterPro; IPRO02347; Adh_short_C2.	
DR	Pfam; PF00106; adh_short; 1.	
DR	Pfam; PF00678; adh_short_C2; 1.	
DR	PRINTS; PR00081; GDHRDH.	
DR	PRINTS; PS00061; ADH_SHORT; 1.	
DR	PROSITE; PS00061; ADH_SHORT; 1.	
KW	Oxidoreductase; NADP; 3D-structure.	
FT	NP_BIND 25 49	NADP (BY SIMILARITY).
FT	ACT_SITE 171 171	BY SIMILARITY.

SQ SEQUENCE 273 AA; 29617 MW; 39A523EF04EA81F1 CRC64;

Query Match 25.5%; Score 357; DB 1; Length 273;
 Best Local Similarity 33.6%; Pred. No. 1.4e-19;
 Matches 89; Conservative 58; Mismatches 110; Indels 8; Gaps 4;

QY 15 NSVRMASSGMRTRDPLANKVALVLTASTDGGFAIARRLAODGAHVVSRRKQNVQDAVA 74
 DB 4 SKVSMNCCNNGRSLKGTALTALVGGGIGVAVIEELAGLAGVYTCRSNEKELDECLE 63

QY 75 TLQGEGLSVGTGVCHVGAEDRERLVAATVAVKLHGG- IDILVSNAAVNPFFGSGIMDVTEEV 133
 DB 64 IWRKGLNVEGSCDLLSRTEDKLMQVAHVDFDKLNLVYNAGV-VIHKAKDFTKED 122

QY 134 WDKTLDINVKPALMTKAVVPEMERKGGGVVIVSSIAAFSPGSPYNNVSKTALLGLT 193
 DB 123 YNIIIMGTNEFAAYHLSQIAYPLLKASQNGNIFLSSIAAGFSAIPSVLSYASKGAINQMT 182

QY 194 KTLATELAPRINRVNCLAPG-----LITKTSFMRMLMDKKEESKMETLRIRRLGEPEDC 248
 DB 183 KSLACAEAKDNIRVNSVAPGVLITPLVETAIKKNPHQKEEIDNFIVKT-PMGRAGKQEV 241

QY 249 AGIVSFLCSEDASYITGETVVVGGG 273
 DB 242 SALIAFLCPRAASYITGOIWDGG 266

RESULT 12

ID_FABG_RICPR STANDARD; PRT; 241 AA.

AC P50941;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
 DE ACYL CARRIER PROTEIN REDUCTASE)
 GN FABG OR RP762.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RL mitochondria."
 RL Nature 396:133-140(1998).
 RN [2]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=94179117; PubMed=8132476;
 RA Dunkin S.M., Winkler H.H., Wood D.O.;
 RT "Isolation and characterization of the Rickettsia prowazekii reca
 RL gene."
 RL J. Bacteriol. 176:1777-1781(1994).
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=96259307; PubMed=8662004;
 RA Andersson S.G.E., Sharp P.M.;
 RT "Codon usage and base composition in Rickettsia prowazekii."
 RL J. Mol. Evol. 42:525-536(1996).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
 CC NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

EMBL; AJ235273; CAAL15190.1;
 DR EMBL; U01959; ; NOT_ANNOTATED_CDS.
 DR HSP; Q12634; 1YBV.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PRINTS; PR00081; GDHRDH.
 DR PRINTS; PR01167; INSADHFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 11 35 NADP (BY SIMILARITY).
 FT ACT_SITE 148 148 BY SIMILARITY.
 SQ SEQUENCE 241 AA; 25759 MW; E43B8711545B8295 CRC64;

Query Match 25.3%; Score 354; DB 1; Length 241;
 Best Local Similarity 33.6%; Pred. No. 1.9e-19;
 Matches 82; Conservative 61; Mismatches 91; Indels 10; Gaps 4;

QY 30 LANKVALVASTDGGFAIARRLAODGAHVVSRRKQNVQDAVATLQGEGLSVGTGVCH 89
 DB 4 LTGKTSITGASSGIGSAIARLLHKLKSGVITSGNEEKLKSLGNALKD---NYTIEVCN 60

QY 90 VKAEDRELRVATVAVKLHGGDILVSNAAVNPFFGSGIMDVTEEVWDKTLDINVKAPALMT 149
 DB 61 LANKKECSNLISKT----SNLDILVCNAGITSDTAIR-MKDQDFKVIDINLKNFILN 115

QY 150 KAVVPEMERKGGGVVIVSSIAAFSPGSPYNNVSKTALLGLTKTLATELAPRINRVNC 209
 DB 116 REAKKMKLOKRYGRIINISSIVGAGNPGQANYCASKAGLIGMTKLSYEVAITGVNA 175

QY 210 LAPGLIKTSFMRMLMDKKEESKMETLRIRRLGEPEDCAGIVSFLCSEDASYITGETVV 269
 DB 176 VAPGFIKSDMTDKL--NEKQREAIQKIPGLTYGIPEDVAVAVAFSLASNNASYITQTILH 233

QY 270 VGGG 273
 DB 234 VNGG 237

RESULT 13

Y4LA_RHISN STANDARD; PRT; 278 AA.

ID Y4LA_RHISN
 AC P55541;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE Y4LA (EC 1.1.1.100).
 GN Y4LA.
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid sym pNGR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes."
 RL Nature 387:394-401(1997).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000082; AAB91754.1; .
 CC HSSP; P29132; 1DFI.
 CC DR InterPro: IPR002198; Adh_short.
 CC DR InterPro: IPR002347; Adh_short_C2.
 CC DR Pfam; PF00106; adh_short; 1.
 CC DR Pfam; PF00678; adh_short_C2; 1.
 CC DR PRINTS; PR00080; SDRFAMILY.
 CC DR PRINTS; PR00081; GDRDH.
 CC DR PROSITE; PS00061; ADH_SHORT; 1.
 CC DR Hypothetical protein; Oxidoreductase; plasmid.
 CC FT NP_BIND 10 34
 CC FT ACT_SITE 159 159
 CC FT ACT_SITE 278 AA; 128743 MW; 1D0105625B9DF2B CRC64;
 CC SQ SEQUENCE
 CC -----

[illegible]

RESULT	14	Y019.THEMA	STANDARD;	PRT;	256 AA.
ID	Y019.THEMA				
AC	Q56318;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	PUTATIVE OXIDOREDUCTASE, TM0019 (EC 1.1.1.1)				
GN	TM0019.				
OS	Thermotoga maritima.				
OC	Bacteria; Thermotogales; Thermotoga.				
OX	NCBI_TaxID=2336;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MSB8 / DSM 3109;				
RX	MEDLINE=96125254; PubMed=8550425;				
RA	Kletzian A., Adams M.;				
RT	"Molecular and phylogenetic characterization of pyruvate and 2-				
RT	ketoisovalerate ferredoxin oxidoreductases from Pyrococcus furiosus				
RL	and pyruvate ferredoxin oxidoreductase from Thermotoga maritima.";				
RL	J. Bacteriol. 178:248-257(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MSB8 / DSM 3109;				

```

MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999)
CC
CC -i- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC between the Swiss Institute of Bioinformatics and the EMBL outstat-
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for commer-
CC entities requires a license agreement (See http://www.isb-sib.ch/annou-
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X85171; CAA59459.1; -
DR EMBL; AE001690; AAD35113.1; -
DR HSSP; P19992; 2HSD.
DR TIGR; TW0019; -
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; P00080; SDRFAMILY.
DR PRINTS; P00081; GDRDHD.
DR PRINTS; P01167; INSADHFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 9 33 NADP (BY SIMILARITY).
FT ACT_SITE 153 153
FT CONFLICT 130 136 BY SIMILARITY.
FT SEQUENCE 256 AA; 28078 MW; D68160B1D7980C6B CRC64.
SQ

```

Query Match 24.9%; Score 349; DB 1; Length 256;
Best Local Similarity 35.2%; Pred. No. 4.9e-19;
Matches 89; Conservative 45; Mismatches 103; Indels 16; Gaps 5;

Qy	30	LANKVALVTASTD	IGFATARRIAQGAHV	VVSSRQQNVQAVAT	LQOEGLSVTGT	VCH	89.	
Db	2	LEGVAVVTGGG	IGAAIAQLFAENG	MKVIAIEIDEEAG	VEREEMLRERGLD	VTFKTD	61	
Qy	90	VGKAEDRERL	VATAVKLHG	GIDILVSNAA	PNPFGSIMD	VTREVPDKTLD	INVKAPALMT 149	
Db	62	VADENSVKN	VRKTVELY	GGVDVLVNA	AVMS-VKSI	FEPLEEN	ERVTRNUTGPICS 120	
Qy	150	KAVVPEMEK	RGGSVVIS	IAAFSPG	SPVNSK	TALLG	LTKTLATLAPRNVNC 209	
Db	121	RYCAEEMIK	RGGVII	INIASTR	AFSEPDE	PEPYSASK	GGLVALTHSLAVLSRYHIRVVS 180	
Qy	210	LAPGLIKTS	PSRMLMD	KKKEES	MTLR-----	IRLPE	PCAGIVFLC-SEDA 260	
Db	181	ISFCWI	TS-----	EW-KKSL	RKKPD	LRPIED	HEQHPAGRVGNPLDIAHL	CVFLADDEKA 234
Qy	261	SYITGETV	VVVGGG	273				
Db	235	GFITG	NFIVDGG	247				

RESULT	15.	
DHG2L_BACSU		
ID	DHG2L_BACSU	STANDARD;
-AC	P80869;	PRT; 258 AA.
DT	01-NOV-1997	(Rel. 35; Created)
DT	15-JUL-1999	(Rel. 38; Last sequence update)
DT	20-AUG-2001	(Rel. 40; Last annotation update)
DE	GLUCOSE 1-DEHYDROGENASE II (EC 1.1.1.47)	(GDPH-II) (GENERAL)

```

Db      185 NAIAPGTIATESN-----VDTKKEESROKLKIPMKAFGKPVEEVAAMLVSEASYVT   240
Qy      265 GETVVVGGGT---PSRL 278
        |||: ||| ||| |||
Db      241 GATLFVDGGMTLFPSQL 257

```

Search completed: January 22, 2002, 15:22:36
Job time: 191 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:19:10 ; Search time 23.8 Seconds
(without alignments)
1708.559 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGGGTPSRL 278

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_17.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1399	100.0	278	4 Q9BT22	Q9bt22 homo sapien
2	1395	99.7	278	4 Q9NV08	Q9nv08 homo sapien
3	1301	93.0	260	4 Q9H3N5	Q9h3n5 homo sapien
4	1293	92.4	260	4 Q9S162	Q9s162 homo sapien
5	1123	80.3	260	6 Q9GKX2	Q9gkx2 oryctolagus
6	1086	77.6	260	14 Q9EQU4	Q9equ4 mus musculus
7	1086	77.6	260	11 Q9S9L2	Q9s9l2 mus musculus
8	824.5	58.9	280	4 Q9H2R2	Q9h2r2 homo sapien
9	787	56.3	282	11 Q9D3M7	Q9d3m7 mus musculus
10	716	51.2	216	11 Q9D2B3	Q9d2b3 mus musculus
11	664	47.5	317	5 Q9VRJ4	Q9vrj4 drosophila
12	636	45.5	257	5 Q16619	Q16619 caenorhabdi
13	614	43.9	260	5 Q93790	Q93790 caenorhabdi
14	588.5	42.1	254	10 Q9S9W2	Q9s9w2 arabidopsis
15	468	33.5	241	11 Q9DC11	Q9dc11 mus musculus
16	459.5	32.8	260	2 Q33308	Q33308 mycobacteri
17	454	32.5	254	2 Q9A3X5	Q9a3x5 caulobacter
18	409	29.2	255	2 Q9I2R7	Q9i2r7 pseudomonas
19	389.5	27.8	253	2 Q9K3Y8	Q9k3y8 streptomyce

20	388	27.7	246	2 Q9KA03	Q9ka03 bacillus ha
21	384.5	27.5	273	2 Q9F5J1	Q9f5j1 streptomyce
22	381.5	27.3	266	2 Q9EWT4	Q9ewt4 streptomyce
23	381.5	27.3	282	1 Q9S564	Q9s564 pyrococcus
24	381	27.2	255	1 Q9HQ41	Q9hq41 halobacteri
25	377	26.9	264	10 Q9LHT0	Q9lht0 arabidopsis
26	376	26.9	261	2 Q54812	Q54812 streptomyce
27	374	26.7	243	2 Q9CHF7	Q9chf7 lactococcus
28	371	26.5	260	2 Q9A8R5	Q9a8r5 caulobacter
29	371	26.5	263	1 Q9V212	Q9v212 pyrococcus
30	370	26.4	249	2 Q9KEB5	Q9keb5 bacillus ha
31	369.5	26.4	246	2 Q9A7P5	Q9a7p5 caulobacter
32	367	26.2	240	2 Q9X429	Q9x429 streptococc
33	366	26.2	261	2 Q54280	Q54280 saccharopol
34	365.5	26.1	248	2 Q31680	Q31680 bacillus su
35	364	26.0	254	2 Q9RT26	Q9rt26 deinococcus
36	363	25.9	274	10 Q96457	Q96457 hyocycamus
37	360.5	25.8	261	2 Q54176	Q54176 streptomyce
38	359	25.7	248	2 Q9KJF1	Q9kjf1 thauera aro
39	358.5	25.6	256	2 P95273	P95273 mycobacteri
40	357.5	25.6	259	2 Q9S3U5	Q9s3u5 bacteroides
41	357.5	25.6	268	1 Q9HLN6	Q9hln6 thermoplas
42	357	25.5	248	2 Q9JXR1	Q9jxr1 neisseria m
43	357	25.5	248	2 Q9JW61	Q9jw61 neisseria m
44	356	25.4	264	10 Q9AR59	Q9ar59 solanum tub
45	354.5	25.3	262	2 Q9L9F8	Q9l9f8 streptomyce

ALIGNMENTS

RESULT 1
Q9BT22 PRELIMINARY; PRT; 278 AA.
ID Q9BT22
AC Q9BT22
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003019; AAH03019.1;
SQ SEQUENCE 278 AA; 29537 MW; 3B06A229E1BBE47B CRC64;

Query Match	100.0%;	Score 1399;	DB 4;	Length 278;
Best Local Similarity	100.0%;	Pred. No. 1.8e-96;		
Matches 278;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	MHKAGLLGLCARAWNSVRMASSGTMTRRDPLANKVALVTASTDGIGFAIARRLAODGAHV 60			
Db 1	MHKAGLLGLCARAWNSVRMASSGTMTRRDPLANKVALVTASTDGIGFAIARRLAODGAHV 60			
Qy 61	VSSRKQNVDAQVATLQGEGLSVTGVCVHGKAEADRLRVATAVKLHGIDILVSNAAVN 120			
Db 61	VSSRKQNVDAQVATLQGEGLSVTGVCVHGKAEADRLRVATAVKLHGIDILVSNAAVN 120			
Qy 121	PFPGSIMDYTEEVWTKTLDINVKAPALMTKAVVPENMEKGGGWSWIVSSIAAFSPSPGFS 180			
Db 121	PFPGSIMDYTEEVWTKTLDINVKAPALMTKAVVPENMEKGGGWSWIVSSIAAFSPSPGFS 180			
Qy 181	PYNVSKTALLGLTKTIAELAPNIRVNCLAPGLIKTISRMLWMDKEESKKTILRIR 240			
Db 181	PYNVSKTALLGLTKTIAELAPNIRVNCLAPGLIKTISRMLWMDKEESKKTILRIR 240			
Qy 241	RLGPEPCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278			

```

Db 241 RLGEPEDCAGIVSLCSEDASITGETVVVGGTSPRL 278
|||||
RESULT 2
Q9NV08 PRELIMINARY; PRT: 278 AA.
AC Q9NV08;
DT 01-OCT-2000 (TrEMBLrel.15, Created)
DT 01-OCT-2000 (TrEMBLrel.15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE CDNA FLJ11008 F15, CLONE PLACE1003100, MODERATELY SIMILAR TO HEP27
DE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Negahari K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
EMBL: AK001870; BAA1953.1; -.
DR InterPro: IPR002198; Adh_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00678; adh_short_C2; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 278 AA; 29509 MW; 6B0096BBA89152A0 CRC64;

Query Match 99.78%; Score 1395; DB 4; Length 278;
Best Local Similarity 99.6%; Pred. No. 3.6e-96;
Matches 277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHKAGLLGLCARAVNSVRMAGSGTTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 60
DB 1 MHKAGLLGLCARAVNSVRMAGSGTTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 60
QY 61 VSRKQNVQAVATLQGEGLSVTGCHVGAEDRERLAVATAVKLHGIDILVSNAAVN 120
DB 61 VSRKQNVQAVATLQGEGLSVTGCHVGAEDRERLAVATAVKLHGIDILVSNAAVN 120
QY 121 PFGSINDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFS 180
DB 121 PFGSINDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFS 180
QY 181 PYNYSKALLGLTKTLATEAPRNIRVNCPLAPGLIKTSFRLMWMKKEESMKETLRIR 240
DB 181 PYNYSKALLGLTKTLATEAPRNIRVNCPLAPGLIKTSFRLMWMKKEESMKETLRIR 240
QY 241 RLGEPEDCAGIVSLCSEDASITGETVVVGGTSPRL 278
DB 241 RLGEPEDCAGIVSLCSEDASITGETVVVGGTSPRL 278

RESULT 3
Q9H3N5 PRELIMINARY; PRT: 260 AA.
AC Q9H3N5;
DT 01-MAR-2001 (TrEMBLrel.16, Created)
DT 01-MAR-2001 (TrEMBLrel.16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.

```

```

GN HUMNRDR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
RT "cDNA cloning and characterization of peroxisomal short-chain
RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
EMBL: AB045131; BAB18775.1; -.
DR InterPro: IPR002198; Adh_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00678; adh_short_C2; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 260 AA; 27572 MW; 0BD2E0C0D2E37D08 CRC64;

Query Match 93.08%; Score 1301; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.2e-89;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 MASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVSVSRKQNVQAVATLQ 78
DB 1 MASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVSVSRKQNVQAVATLQ 60
QY 79 EGLSVTGTVCHVGAEDRERLAVATAVKLHGIDILVSNAAVNPFSGIMDVTEVWDKTL 138
DB 61 EGLSVTGTVCHVGAEDRERLAVATAVKLHGIDILVSNAAVNPFSGIMDVTEVWDKTL 120
QY 139 DINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFSVYNVSKTALLGLTKTLAI 198
DB 121 DINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFSVYNVSKTALLGLTKTLAI 180
QY 199 ELAPRNIRVNCPLAPGLIKTSFRLMWMKKEESMKETLRIRRLGEPEDCAGIVSLCSE 258
DB 181 ELAPRNIRVNCPLAPGLIKTSFRLMWMKKEESMKETLRIRRLGEPEDCAGIVSLCSE 240
QY 259 DASVITGETVVVGGTSPRL 278
DB 241 DASVITGETVVVGGTSPRL 260

RESULT 4
Q95162 PRELIMINARY; PRT: 260 AA.
AC Q95162;
DT 01-MAY-1999 (TrEMBLrel.10, Created)
DT 01-MAY-1999 (TrEMBLrel.10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
GN SCAD-SRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fransen M., Van Veldhoven P.P., Subramani S.;
RT "Identification of two novel mammalian proteins with a C-terminal
RT peroxisomal targeting signal by using the pVI-phase display
RT technology.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
EMBL: AF044127; AAD02292.1; -.

```


DR HSSP: P50163; 2AE1
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27602 MW; 83687CACD82B9B5 CRC64;

Query Match 92.4%; Score 1293; DB 4; Length 260;
 Best Local Similarity 99.2%; Pred. No. 1.3e-88;
 Matches 258; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 60

Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 120

Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 180

Qy 199 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

Qy 259 DASYYTGTGVVGGTSPSL 278
 Db 241 DASYYTGTGVVGGTSPSL 260

RESULT 5
 Q9GKX2 PRELIMINARY; PRT; 260 AA.
 AC Q9GKX2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.
 GN RABNROR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
 RT "cDNA cloning and characterization of peroxisomal short-chain
 RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
 CC FAMILY.
 DR EMBL; AB045133; BAB18776.1;
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27429 MW; 5B0585B58911B90C CRC64;

Query Match 80.3%; Score 1123; DB 6; Length 260;
 Best Local Similarity 85.4%; Pred. No. 5.6e-76;
 Matches 222; Conservative 18; Mismatches 20; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 60

Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 120

Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 180

Qy 199 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

Qy 259 DASYYTGTGVVGGTSPSL 278
 Db 241 DASYYTGTGVVGGTSPSL 260

RESULT 6
 Q9EQU4 PRELIMINARY; PRT; 260 AA.
 AC Q9EQU4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.
 GN MOUNROR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57 BLAK; TISSUE=LIVER;
 RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
 RT "cDNA cloning and characterization of peroxisomal short-chain
 RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
 CC FAMILY.
 DR EMBL; AB045133; BAB18776.1;
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR003015; HLH_Myc.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDFAMILY.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27726 MW; 005C5F19AC539CF6 CRC64;

Query Match 77.6%; Score 1086; DB 11; Length 260;
 Best Local Similarity 81.5%; Pred. No. 3.1e-73;
 Matches 212; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVASTDGIQFAIARRLAQDGAHVSVSSRKQONVDQAVATLQ 60

Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGIDILVSNAAVNPFGSIMDVTEEVDKTL 120

Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYNSKVTALLGLTKTLAI 180

Qy 199 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRINVCNAPGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

```

Db      181 ELAPKNIRVNCPLAIKTRFSSVLWEKEAREDFIKEAMQIRRLGKPEDCAGIVSFLCSE 240
Qy      259 DASYITGETVVVGGSPTSLR 278
        ||||| ||||||| |||||||
Db      241 DASYINGETVVVGGSPTSLR 260
        ||||| ||||||| |||||||

RESULT  7
ID Q99LB2 PRELIMINARY; PRT; 260 AA..
AC Q99LB2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SIMILAR TO PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR ENBL; BC003484; AAH03484.1;
SQ SEQUENCE 260 AA; 27754 MW; 021733159D2BC3C9 CRC64;

Query Match 77.6%; Score 1086; DB[1]; Length 260;
Best Local Similarity 81.5%; Pred. No. 3.le-73;
Matches 212; Conservative 22; Mismatches 26; Indels 0; Gaps

Qy      19 MASSGTRRDPLANKVALVTASTDGIGFAIARLAODGAHVYVSSRKQONVDQAVATLOG 78
        ||||| :|||: |||: |||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MASSGLTRRNLNKNVALVTASTDGIGFAIARLAEDGAHVYVSSRKQONVDRAVATLOG 60
        ||||| :|||: |||: |||: ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      79 EGLSVGTCTCHGVKAEDRELRYATAYKLHGIGDILYSNAAVNPFFGISMDVTEEVMDKTL 138
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 EGLSVGTICVHGKAEDEKLIITALKRHGGIDILYSNAAVNPFFGNLMDVTEEVMDKVL 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      139 DINVKPALMTKAVPEMEKRGGCVVIYSSSTAASFSPGFSPYNVSKTALLGLTKTLAI 198
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 SINVTAMMIKAVPEMEKRGGCVVIYVSGVAGTFRFPSLGPYNVSKTALLGLTKNEAA 180
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      199 ELAPNRIRVNCLAPGLIKTSFSRLMWMDEKESMKETLRIRRLGEPCAGIVSFLCSE 258
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 ELAPKNIRVNCPLAIGLIKTFSSVLWEKEAREDFIKEAMQIRRLGKPEDCAGIVSFLCSE 240
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      259 DASYITGETVVVGGSPTSLR 278
        ||||| ||||||| |||||||
Db      241 DASYINGETVVVGGSPTSLR 260
        ||||| ||||||| |||||||

RESULT  8
Q9H2R2 PRELIMINARY; PRT; 280 AA..
ID Q9H2R2;
AC Q9H2R2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SHORT-CHAIN ALCOHOL DEHYDROGENASE.
GN HEP27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC "Human Hep27 chromosomal gene";
RA Pellegrini S., Censini S., Guidotti S., Covacci A., Gabrielli F.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
```

RESULT	10
Q9D2U3	
ID	PRELIMINARY; .PRT; 216 AA.
AC	Q9D2U3
AD	Q1-JUN-2001 (TReMBLrel_17; Created)
DT	Q1-JUN-2001 (TReMBLrel_17; Last sequence update)
DDT	Q1-JUN-2001 (TReMBLrel_17; Last annotation update)
DE	DNA SEGMENT, CHR 14, UNIVERSITY OF CALIFORNIA AT LOS ANGELES 2.
DN	D14UCLA2.
GN	OS
OS	Mus musculus (Mouse).
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OCC	Mammalia; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
[1]	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6J; TISSUE=CEREBELLUM;
RX	MEDLINE=21085660; PubMed=12117851;
RC	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahara I.,
RA	Saito T., Okazaki Y., Gojohji T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schiml L.M., Staubli F., Suzuki R., Tomita M., Wegner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Anno H., Baldarelli R., Barsh G.,
RA	Blake J., Bellodi D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.

RESULT 11.

ID	Q9VRJ4	PRELIMINARY:	PRT;	317 AA.
AC	Q9VRJ4;			
DT	01-WAY-2000 (TREMBLrel. 13, Created)			
DT	01-WAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CG10672 PROTEIN.			
GN	CG10672.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscophora;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
ON	NCBI_Taxid=7227;			
RX	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=BERKELEY;			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Adams M.D., Celisnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.			
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Baltes R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delecher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.			

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; AF003567; AAF50801.1; -.
DR HSP; P50162; IAE1.
DR FlyBase; FBgn0035588; CG10672.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; adh_short_C2.
DR Pfam; PF00106; adh_short_1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
DR Oxidoreductase.
SQ SEQUENCE 317 AA; 33617 MW; 0D8C30A7E64EE4E3 CRC64;

Query Match 47.5%; Score 664; DB 5; Length 317;
Best Local Similarity 50.9%; Pred. No. 1e-41;
Matches 136; Conservative 50; Mismatches 79; Indels 2; Gaps 1;

QY 12 RAWNSVMSASGWTTRPLANKVALVTASTDIGIGFAIARRLAQDAHVVSRRKQNVQD 71
DB 53 RLSSSSOSTAGTKMKR--LAGKAVAVTASTDIGIFAIKRAEDGAAVVSSRRKQNVDS 110
QY 72 AVATLOGEGSLVTCVGHVGAEDRERLAVATKVLHGIDILVSNAAVNPFFGSIIMDYTE 131
DB 111 ALAELRLNVLNVHGLKCHVSEPDQRKQLFEETISKFKNLNVLVSNAAVNPFFGSLVCEDE 170
QY 132 EWDKTLIDINVKAPALMKAVPPEMEKRGSGSVIVSVISIAAFSPGSPYNSKTLALIG 191
DB 171 KWDXKIFDVNVKSSYLAKAELPLRQKNSSIVFVSIAGYDAFELLGAYSVSKTALIG 230
QY 192 LTKTLALAPRNTRVNCPLAPGLTKTSFRMLMDKEESMKETLIRIRLGEPEDCAGI. 251
DB 231 LTRAAAKDLAPEGIRVNCPLAPGVIRTFKSKALYENESANEALSKIPMGRIGTSEENAGV 290
QY 252 VSELCSEDAVIITGVVGGTSPRL.278
DB 291 VSEFLVSDAGIYITGESIVAGGGMATRL 317

RESULT 12
ID Q16619 PRELIMINARY; PRT; 257 AA.
AC Q16619
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE F36H9.3 PROTEIN.
GN F36H9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;

```

```

RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Consey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RN Nature 368:32-38(1994).
RL [2]
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Dante M., Kramer J., Twyman B.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; AF016668; AAB66092.1; -.
DR HSP; P50162; IAE1.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short_1.
DR Pfam; PF00678; adh_short_C2; 1.
DR Oxidoreductase.
SQ SEQUENCE 257 AA; 27617 MW; FBF35DE9DEBEEIF CRC64;

Query Match 45.5%; Score 636; DB 5; Length 257;
Best Local Similarity 50.6%; Pred. No. 9.3e-40;
Matches 126; Conservative 49; Mismatches 74; Indels 0; Gaps 0;

QY 30 LANKVALVTASTDIGIGFAIARRLAQDAHVVSRRKQNVQDAVATLOGEGSLVGTVCH 89
DB 9 LTRVALVTASTDIGIGFAIARRLAQDAHVVSRRKQNVQDAVATLOGEGSLVGTVCH 68
QY 90 VGKAEDRERLAVATKVLHGIDILVSNAAVNPFFGSIIMDYTEVWDKTLIDINVKAPALMT 149
DB 69 VGKSDRERLAVATKVLHGIDILVSNAAVNPFFGSIIMDYTEVWDKTLIDINVKAPALMT 128
QY 150 KAVVPEMEKRGSGSVIVSVISIAAFSPGSPYNSKTLALIGLTKLALAPRNIRVNC 209
DB 129 KEAVPHEASGRGNVVFSSVAGYSPMNEICAYSVMKTLTGLSKSLALNARNIRVNS 188
QY 210 LAPGLIKTSFRMLMDKEESMKETLIRIRLGEPEDCAGIVFLCSEDAVIITGVTV 269
DB 189 IAPGLIKTSFRMLMDKEESMKETLIRIRLGEPEDCAGIVFLCSEDAVIITGVTV 248
QY 270 VGGTSPRL.278
DB 249 INGMHARI 257

RESULT 13
Q93790 PRELIMINARY; PRT; 260 AA.
ID Q93790
AC Q93790
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE F54F3.4 PROTEIN.
GN F54F3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.

```


Search completed: January 22, 2002, 15:22:21
Job time: 191 sec